THE PROCEEDINGS OF THE PHYSICAL SOCIETY

Section B

VOL. 65, PART 8 1 August 1952 No. 392 B

THERMAL PROPERTIES AND GROWTH OF CRYSTALS

The following six contributions were read at the Society's meeting at Southampton on 19th December 1951.

The Growth from the Melt of Single Crystals of Low Melting Point Metals

BY A. J. GOSS AND S. WEINTROUB
University of Southampton

MS. received 10*th March* 1952 ; *read before the Physical Society at Southampton on* 19*th December* 1951

ABSTRACT. The results of a series of experiments on the preparation of tin, lead, zinc, and bismuth single crystals in the form of rods are given. The effects on the crystallization of various thermal factors, in particular the temperature gradient and the rate of growth, are studied and the results for the four metals are compared.

§ 1. INTRODUCTION

THE study of the physical properties of metals is greatly simplified by the use of specimens in the form of single crystals. Various techniques have been proposed and developed during the past forty years for the growth of such crystals but a close examination of the methods reveals that they are still largely empirical. Holden (1949), to whom we are indebted for the most recent comprehensive review of single crystal techniques, makes it clear that the information available as to the conditions governing the successful growth of metal single crystals is but meagre and very often contradictory. Several investigators, however, consider that the purity of the metal, the temperature gradient G in the specimen, and the rate R of growth of the single crystals, are controlling factors, but no systematic investigation of the influence of these or other factors on the successful growth of metal single crystals has as yet been made. Experiments to provide information on this subject are being carried out in this laboratory and some of the results so far obtained for pure tin*, lead, zinc and bismuth are described and discussed in this paper.

The horizontal travelling furnace method of Andrade and Roscoe (1937) was used and specimens of the four low-melting point metals mentioned above were crystallized, by slow solidification from the melt, into single crystal rods approximately 5 mm in diameter and some 20–30 cm in length. About 100 specimens of each metal were investigated. The temperature gradient in the solidifying metal was varied over as wide a range as possible, and various rates of growth, from 0·06 to 30 mm/min, were used. The variation in the temperature of the metal was recorded by means of a thermocouple embedded in a dummy specimen and the average temperature gradient at the solid–liquid boundary deduced. This is the gradient which is quoted in the tables. The rods, after crystallization, were carefully examined, the crystalline structure determined and the crystallographic orientations of the single crystals measured. The results, together with other relevant details, are listed and discussed in the following sections. Details of the experimental arrangements, theoretical considerations and the influence of varying amounts of added impurities are omitted. It is hoped to describe these later.

---

* The results for tin have been briefly reported in *Nature, Lond.*, 1951, **167**, 349.

§ 2. THE INFLUENCE OF G AND R ON THE FORM OF THE CRYSTALS

It is well known that single crystals are not perfect but frequently contain perceptible imperfections of various kinds which are referred to in the literature by such names as lineage, mosaics, striations, etc. The rods, after crystallization, were first examined visually and the imperfections seen carefully noted. The rods were then chemically etched and a sketch made of the crystalline structure made visible by the etching. When 10 cm or more of a rod was found to consist of one predominant single-crystal structure, that rod was classified as single. The orientation of the cylindrical axis of the rod, relative to the crystallographic axes of the crystals formed in the rod, was determined, by means of a specially designed optical goniometer, by the method used by Chalmers (1935).

For the four metals studied, it was found that, on an average, though exceptions did occur, the crystalline forms of the rods of any one metal which were produced with the same values of G and R followed the same general pattern. This 'average', over 10–20 or more specimens, is the subject of the rows in the tables 1–4. In the columns are listed the number of specimens examined in the particular group, the percentage of single crystals obtained, the percentage with imperfections and the values of G and R used.

(i) *Tin*

In table 1 the results for tin are summarized. The metal was 99·997% pure and was obtained from Messrs. Johnson, Matthey & Co. Ltd. In most cases the first few centimetres of the rod to solidify consisted of small crystals and the remainder of the rod usually formed a single crystal. Varying G had little or no effect on the crystalline form but varying R altered the results considerably. The change of shape of the temperature–time trace at solidification was also altered by the variation in R (see § 4 (i)).

The polycrystals within the first centimetre of the rods were 1 mm² or less in surface area and larger crystallites appeared as the distance from the end of the rod increased. The inter-crystalline boundaries were generally jagged, especially near the beginning of the rod. Experiments performed with a thermocouple embedded within this portion of the specimens indicated that supercooling occurred in this case. For the slowest speeds, the first section of the rod consisted very often of a single crystal but of different orientation from that of the remainder.

For rates of growth of 0·5 mm/min and 20 mm/min, apart from the first few centimetres, the remainder consisted of a single crystal for which the orientation was constant throughout the whole length of the rod. Three main types of imperfections were observed in the single crystal portions: (*a*) small inserts of crystals of different orientations; (*b*) two or more crystals within the length and (*c*) lineage, i.e. thin strips of crystallite, less than 1 mm in width. The crystals in (*b*) often formed a bicrystal with the boundary parallel to the rod.

Table 1. Tin—Tetragonal Structure 99·997% Purity

(1)	(2)	(3)	(4)	(5)
20	16	100	0	11—13
10	20	70	20	8—10
5	24	42	50	9—19
1·5	24	67	13	10—28
0·5	24	96	4	10—35

(1) rate of growth R (mm/min); (2) number of specimens; (3) percentage of single crystals; (4) percentage of crystals with lineage; (5) range of gradient G (°c/cm).

Rates of growth of 5, 1·5 and 0·5 mm/min produced the (*b*) type of imperfection. The strips of crystallite in (*c*) grew along the length of the rod and usually had orientations differing from one another by one or two degrees. In the case of tin, for specimens containing lineage, the angle between the rod axis and the *c*-axis was always 70–90° and the orientations of the individual strips were usually related to one another by small angles of rotation of the *c*-axis about the rod axis. This type of imperfection was greatest for the 5 mm/min rate of growth. Lineage formations in tin have been studied, quite recently, by Teghtsoonian and Chalmers (1951).

Another type of macro-mosaic which occurred, and which was observed both before and after etching, consisted of fine lines about 0·1 mm apart, but no variation in orientation was detected on examination of the surfaces containing these lines.

On crystals grown at the rate of 20 mm/min various types of rapid solidification phenomena, such as dendritic markings on the surface or occasional shrinkage cavities, were seen near the end of the specimens.

(ii) *Lead*

Tadanac lead of 99·998% purity, kindly presented by Messrs. Goodlass Wall and Lead Industries, Ltd., and 99·99% pure lead obtained from Messrs. Johnson, Matthey & Co. Ltd., were studied. No significant difference could be detected between the two in regard to crystallization. The results summarized in table 2 refer only to Tadanac lead.

As for tin, the effect of varying the temperature gradient had negligible effect, but it is just possible that higher gradients might influence the crystallization. In more than half of the 143 specimens examined one type of imperfection, lineage, was observed. The strips constituting the lineage were about 1 mm wide but they were not uniformly parallel to the length of the specimen. The orientations of the various strips were related to one another by rotations of the lattice around the specimen axis irrespective of the actual orientation of the crystal. In columns 4 and 5 of table 2 the percentages of specimens with lineage, where the variation in orientation is greater than 1° or 5° respectively, are given.

It can be seen that the greatest proportion of crystals with lineage occurred for rates of growth of 1·5 and 5 mm/min and for these speeds the number of crystals with a variation in orientation of more than 5° was also a maximum. Only for 31 mm/min were rods crystallized free from lineage, although, even for these, variations in orientation of less than ½° could be detected in some cases.

Table 2. Tadanac Lead—Face-Centred Cubic Structure 99·998% Purity

(1)	(2)	(3)	(4)	(5)	(6)
31	13	100	0	0	11
16·5	36	33	67	8	6—21
5·3	15	20	72	53	23—50
1·5	33	18	79	64	5—65
0·5	6	50	50	50	17
0·2	40	62	38	17	5—50

(1) rate of growth R (mm/min); (2) number of specimens; (3) percentage of single crystals; (4) percentage of crystals with lineage $>1°$; (5) percentage of crystals with lineage $>5°$; (6) range of gradient G (°c/cm).

With rates of growth up to 16 mm/min almost every specimen consisted of 20 cm or more of one predominant orientation, but in other specimens inclusions of crystals of other orientations appeared along the length of the rod. For 31 mm/min there was a tendency to form a series of single crystals of 5–15 cm in length with the crystal boundaries across the rod. For 16 or 31 mm/min rapid solidification patterns were formed similar to those for tin.

An average was taken of the number of seed crystals formed in the first millimetre or so of the specimens and, for rates of less than 5 mm/min, it would appear that more than two crystallites could have acted as seeds and, for higher rates, less than two.

To observe the effect of oxidation several crystals were left for 24 hours in air before etching and coloured oxide films formed on some of the crystal faces. Observations on 15 specimens treated in this way showed no oxidation on the (100) planes. This is in disagreement with Gwathmey *et al.* (1945) who found these planes to be active, though at a higher temperature 150–250° c.

(iii) *Zinc*

Crystals were grown from 99·997% pure zinc kindly presented by the National Smelting Co. Ltd. and also from zinc of 99·99% and 99·95% purity. Table 3 contains results of the experiments on the purest metal. Results for the less pure metal were similar but with a definite inclination for the formation of slightly less successful single crystals. For example, for the 5 mm/min rate, the crystals in some cases consisted entirely of strips and, for 1·5 mm/min, bicrystals were frequent.

The size and form of the crystals changed rapidly with the rate of growth but changing the temperature gradient had little effect. The general picture is shown in table 3, but in particular the following are worthy of note:

(*a*) 16·5 *mm/min*. The beginning of each rod always consisted of polycrystals and then strips of crystal about 1 mm wide which extended along the length of the specimen. The orientations of these strips were not simply related to one another as for lineage in tin and lead.

(*b*) 5·3 *mm/min.* The rods in this case differed from those in (*a*) in that the last 10 cm or so consisted either of a bicrystal with the boundary parallel to the specimen, or of a single crystal.

(*c*) 1·5 *mm/min.* Large strips of crystals approximately 1 mm wide occurred at the beginning of the rod and then for over 80% of the specimens the remainder of the rod was a single crystal, the others being bicrystals.

(*d*) 0·5 *mm/min.* The beginning of the rod consisted of one or two crystals and for 90% of the specimens the remainder of the rod was monocrystalline. Inclusions within this length seldom occurred but the crystal often lost its original uniform cross section with the production of a neck or flat surface, or by filling the mould completely over a short distance.

Table 3. Zinc—Hexagonal Structure 99·997% Purity

(1)	(2)	(3)	(4)	(5)	(6)
16·5	20	0	Strips	70—90	8—30
5·3	27	7	Bicrystals	55—90	2—40
1·5	26	85	Bicrystals & Single	12—89 Average—76	3—45
0·5	30	90	Single Crystals	12—90 Average—60	8—30

(1) rate of growth R (mm/min); (2) number of specimens; (3) percentage of single crystals; (4) crystal form; (5) $\psi°$; (6) range of gradient G (°c/cm).

For examination of the crystals thermal etching was sufficient in most cases. Thermal etching produced a surface with ridges parallel to the basal plane. Crystals with small angles between the hexagonal axis and the rod axis, and which were subjected to prolonged thermal etching, exhibited flat faces with ridges. Hexagonal prism planes of the first and second type were developed and the angle of 30° between them could be measured quite accurately.

(iv) *Bismuth*

The purest bismuth available in sufficient quantity was of 99·97% purity and was supplied by Messrs. Mining and Chemical Products Ltd. The crystals produced were less perfect than those of the other three metals. In fact, no particular rate of growth or temperature gradient resulted in a good yield, although rates ranging from 0·06 to 21 mm/min and gradients in the liquid metal from 8 to 32° c/cm were tried.

As little effect was expected from a change of the temperature gradient only the highest and lowest gradients available were used and these produced no change. The considerably wide range of rates covered is indicated in table 4.

Change of the rate of crystallization had a marked effect on the crystal form. This is considered under the two following headings:

(*a*) 1·5–0·2 *mm/min.* These rates resulted in rods with, usually, one or two main crystals and with small inserts of varying crystal orientation. Of the 33 specimens examined only four had lineage with orientations varying by less than 1°; the others had lineage similar to that in tin and lead. This lineage was generally less continuous than that in tin and of a more jagged appearance but it was invariably along the specimen.

(*b*) 5·3–10·6 *mm/min.* Rods were produced consisting of one or two crystals, but broken up transversely, sometimes by short lengths of crystals of different orientations but frequently by 'block' structure. These 'blocks' consisted of units, a few millimetres or less in length, across the rod and with near orientations so that the same crystal orientation extended approximately along the whole length of the rod. Seventeen of the crystals grown at the above rates showed block type imperfection, the others being single crystals. No lineage was visible in these specimens. Except for one example at 1·5 mm/min the block structure was confined to crystals grown at rates of 5·3–10·6 mm/min.

Table 4. Bismuth—Rhombohedral Structure 99·97% purity

(1)	(2)	(3)	(4)	(5)	(6)	(7)
10·6	12	3	Short Lengths of Crystal or 'Blocks'	71·7	31	15
5·3	12	4	'Blocks' or 'Inserts'	71·6	27	20—30
1·5	12	2	Lineage, some Inserts	71·8	3	15—40
0·5	12	2	Lineage and Inserts	78·7	0	20—50
0·2	9	0	Lineage and Inserts	77·4	0	40

(1) rate of growth R (mm/min); (2) number of specimens; (3) number of single crystals; (4) crystal form; (5) average ψ°; (6) maximum supercooling (°c); (7) range of gradient G (°c/cm).

§ 3. THE CRYSTAL ORIENTATIONS

A specially designed optical goniometer was used to measure the orientations of the various crystals grown. The orientation could be determined to within 1°. The variation in orientation along the rod could be observed directly by moving the rod vertically without rotation. Values of the orientation ψ, i.e. the angle between the axis of the rod and the significant crystallographic axis of the metal discussed, are listed in the appropriate table or mentioned below.

(i) *Tin*

The angle between the rod axis and both the [001] and [010] axes were measured. The results are shown in the stereographic projections (figs. 1 and 2).

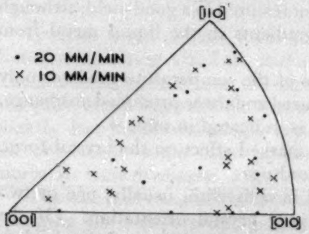


Fig. 1. Stereographic projection of the rod axes of tin crystals grown at fast rates.

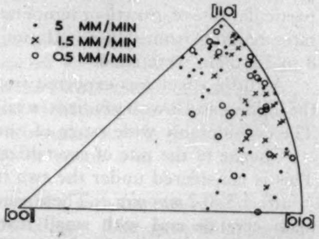


Fig. 2. Stereographic projection of the rod axes of tin crystals grown at medium and slow rates.

For the crystals grown at rates of 10 and 20 mm/min the distribution of the directions of the rod axis relative to the crystallographic axes is approximately uniform but with some preference for the higher values of ψ. For the lower speeds, however, the angle ψ (relative to the [001] axis) is always between 68° and 90° and there is some preference also for the [110] axis to be near to the rod axis. The above values refer to the single crystal portions of the rods. The crystallites in the first few centimetres of the rods had values of ψ usually smaller, but occasionally larger.

(ii) *Lead*

The readings taken on the optical goniometer for the various single crystals of lead were plotted on a Wulff net and the angles between the nearest [100], [110] and [111] axes and the rod axes were determined. A stereographic plot of the directions of the rod axes did not show any particular preferred orientation. The distribution of the orientations was examined statistically, there being values available for more than one hundred specimens. A description of the method used is given elsewhere, and the term 'uniform distribution' found in this way is defined more explicitly by Goss (1951). The examination showed that the distributions for the crystallographic axes were uniform. This applied to both the pure and less pure lead.

(iii) *Zinc*

The main imperfection has been indicated previously as strips of crystal, about 1 mm wide, along the length of the specimen. These strips are not of near orientations as for the lineage specimens of the other metals but consist of crystals with the hexagonal axis approximately at right angles to the rod axis and in all directions around the rod axis.

ψ for the *c*-axis was between 75° and 90° for all the crystals grown at 5·3 mm/min except one for which ψ was 55°. For the fastest rate all the crystals consisted of strips with high values of ψ. The tendency for high values of ψ was still very marked even for the rate of 1·5 mm/min where 78% of the crystals and the bicrystals had ψ's above 75° (table 3).

For the slowest rate, 0·5 mm/min, a large scatter of values of ψ occurred, but there was a slight preference for both the high and low orientations. Low orientation crystals have perfect cleavage planes but the crystals were not usually examined by cleaving. Lineage was observed in a few cases but any lineage at the higher rates was masked by the production of strips instead of single crystals.

(iv) *Bismuth*

The bismuth crystals, both the single and the more usual imperfect crystals, grown at all speeds, favoured one particular orientation. ψ for the [111] axis, the axis of trigonal symmetry, was 60°–90° for 85% of the specimens. In column 5 of table 4 the average values of ψ obtained are given. For a uniform distribution ψ should have an average of 60°, but the average for the rates of 0·5 and 0·2 mm/min is 78°, and for the higher rates 72°. An extremely small, but perhaps significant, effect due to the temperature gradient appears to occur, for the average value of ψ for the crystals grown under a low gradient was at least 5° less than the corresponding value for the crystals grown under a high gradient in each of the rate groups.

Within the first centimetre or so of the rod a crystal with a low value of ψ sometimes occurred, but the other crystallographic axes showed no preference for any particular direction.

.B
5
2

§4. THE TEMPERATURE MEASUREMENT IN THE METALS

Chromel–alumel thermocouples were used for the measurement of the temperature, a silver-soldered couple being immersed in the liquid metal of a dummy specimen identical with the crystal specimens. The e.m.f. was recorded on a Kent potentiometric recorder the scale of which was 250 mm wide measuring to 6 mv. This maximum corresponds to about 150° c. Figure 3 is a reproduction of a typical trace. It represents the variation with time of the temperature of the mid-point of a rod and, for approximately uniform conditions over the centre section of the rod, it gives, in addition, the temperature gradient at any instant along the rod. The temperature gradients were calculated from the traces. The traces were confirmed by independent measurements using a direct reading potentiometer. Further, in order to assess the direction of heat exchange between the specimens and their surroundings measurements were made with differential thermocouples and pairs of thermocouples. The results are of considerable interest for the discussion of the heat flow during crystallization.

The value for the temperature gradient calculated from the curves and quoted in the tables is the average of the gradient in the liquid metal immediately above the melting point and of the gradient in the solid just below the melting point. This value is not an indication of the shape of the cooling curve.

(i) *Tin*

Change of the rate of growth produced a distinctive change in the corresponding traces. Three types of trace were obtained: (i) For the rate of 10 mm/min or more, the temperature at solidification was steady at 232° c, in some cases for a period of minutes, and then fell sharply. The solid gradient was greater than the liquid gradient (fig. 3). (ii) For the rate of 5 mm/min the cooling curve was smooth with no appreciable kink. (iii) For low rates there was a small kink at the melting point and the solid gradient was less than the liquid gradient in contrast to (i) (fig. 3).

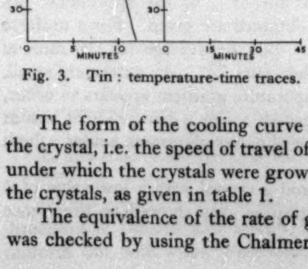


Fig. 3. Tin : temperature-time traces.

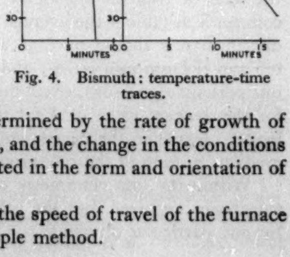


Fig. 4. Bismuth : temperature-time traces.

The form of the cooling curve is thus determined by the rate of growth of the crystal, i.e. the speed of travel of the furnace, and the change in the conditions under which the crystals were grown was reflected in the form and orientation of the crystals, as given in table 1.

The equivalence of the rate of growth and the speed of travel of the furnace was checked by using the Chalmers (1949) ripple method.

(ii) *Lead*

The temperature records were similar to those for tin. For the slow rates of 0·2, 0·5 and 1·5 mm/min there was often a decrease in the temperature gradient as the metal cooled through the temperature of 327° c (the melting point), the effect being more marked at the higher gradients. Unlike tin there was no small kink in the trace at low rates. For rates of 16·5 and 31 mm/min the temperature was approximately constant at or below the solidification point, with the gradient after solidification greater than that immediately before solidification. This behaviour is similar to that of tin.

The form of the cooling curve corresponded directly with the movement of the furnace and confirmation of the results was obtained by using thermocouples connected to a direct reading potentiometer and by using ripples.

(iii) *Zinc*

The chromel–alumel thermocouples were attacked by the zinc vapour, but by using a fresh thermocouple for each run the temperature in the solidifying metal was recorded. At the same time comparison readings were taken on a similar thermocouple shielded from the vapour but still within the furnace tube. For the actual crystallization runs, records were taken of the shielded thermocouple only and the temperature of the solidifying metal was deduced.

For the highest rate of 16·5 mm/min there was an increase in the temperature gradient after solidification. A slight increase in the gradient occasionally occurred at lower rates, but no decrease in gradient was ever observed for low rates of crystallization. This is in contrast to the results obtained with tin and lead.

At low rates and gradients the temperatures of the shielded thermocouple and of the metal corresponded closely, but at high rates and gradients the temperature gradient in the furnace tube was considerably greater than that in the metal. In an extreme case there was as much as 100° difference in temperature between the metal and its surroundings.

The gradient values in table 3 are those deduced for the metal.

(iv) *Bismuth*

Although bismuth slowly attacked the thermocouple wires, measurement of the temperature in the solidifying metal was still possible and records were obtained in the same way as for tin and lead.

Supercooling was observed in almost all the records for rates of 5·3 mm/min and higher, but at slow rates only one instance of supercooling was detected. However, in several cases, a change in gradient for these slow rates took place at 10°–20° c below the solidification temperature.

Bismuth has a much smaller thermal conductivity in the solid than in the liquid ($K_{\text{solid}}/K_{\text{liquid}} = 1/2$). Hence the slope of the temperature–time trace increases sharply as the temperature drops through the solidification point. This occurred regularly except for the lowest rates of 0·2 and 0·06 mm/min. No change in gradient was detected in those cases. These effects are shown in fig. 4.

§ 5. CONCLUSION

The metals tin and lead were similar in their crystallization. Good crystals were produced for both fast and slow rates of growth but, for medium rates, the crystals were less satisfactory and had the greatest amount of lineage.

For bismuth, all rates produced single crystals though many had slight imperfections. Fast rates tended to break up the crystals into blocks, low rates to produce lineage. For this metal, at rates of 5 or more mm/min, supercooling occurred, and this could be seen clearly on the temperature–time traces. This supercooling was peculiar to bismuth.

Zinc behaved quite differently from the other three metals. Only at slow rates were single crystals produced. At faster rates strips of crystal of similar orientations appeared.

In no case was the magnitude of the temperature gradient an influence on the growth, but the form of the temperature–time trace was dependent (in zinc only to a small extent) on the rate of growth.

The rate has, in addition, an effect on the orientation. This is the case for tin and zinc and to a slight extent also for bismuth, but the influence varies from metal to metal. The effects on the various metals together with a summary of the other results are given in table 5.

Table 5. Summary

Rate (mm/min)		Tin	Lead	Zinc	Bismuth
>5	Crystal form	Single crystal	Single crystal	No single crystals	Single crystals with imperfection
	Orientation	Random	Random	ψ tending to 90°	ψ tending to 90°
	Temp. grad.*	Zero at m.p.: increase in liquid	Zero at m.p.: increase in liquid	Increase in solid	Supercooling: increase in solid
5	Crystal form	Lineage	Single crystal with lineage	Bicrystal	Single crystals with imperfection
	Orientation	$70^\circ<\psi<90^\circ$	Random	$55^\circ<\psi<90^\circ$	ψ tending to 90°
	Temp. grad.*	Little change	Little change	No change observed	Supercooling: Increase in solid
<5	Crystal form	Single crystal	Single crystal	Single crystal	Single crystals with lineage
	Orientation	$70^\circ<\psi<90^\circ$	Random	Random	ψ tending to 90°
	Temp. grad.*	Decrease in solid	Decrease in solid	No decrease in solid observed	Increase in solid

* At solidification.

Though the matter is not discussed in this paper it can be shown that the information given by the temperature–time curves for tin, lead and bismuth may be related theoretically to the heat flow from the liquid to the solid by taking into account the release of latent heat, and a value for the ratio of the thermal conductivities of the solid and liquid metals may then be deduced.

The importance of the influence of the purity of the metal on the successful growth of single crystals should not be overlooked, and the results of a further series of experiments in which the effects of added impurities to tin have been investigated will be reported elsewhere. The effect observed with the less pure zinc mentioned earlier appears to fall into line with these subsequent observations.

ACKNOWLEDGMENTS

The authors are indebted to Professor A. M. Taylor and the College authorities for facilities provided, and one of them (A. J. G.) to the Ministry of Education for a Further Education and Training Grant.

REFERENCES

ANDRADE, E. N. DA C., and ROSCOE, R., 1937, *Proc. Phys. Soc.*, **49**, 152.
CHALMERS, B., 1935, *Proc. Phys. Soc.*, **47**, 733 ; 1949, *Proc. Roy. Soc.* A, **196**, 64.
GWATHMEY, A. T., LEIDHEISER, H., and SMITH, G. P., 1945, *National Advisory Committee for Aeronautics*, Technical Note 982.
GOSS, A. J., 1951, *Research*, **4**, 292.
HOLDEN, A. N., 1949, *Amer. Soc. Met.*, Preprint No. 35.
TEGHTSOONIAN, E., and CHALMERS, B., 1951, *Canad. J. Phys.*, **29**, 370.

Mr. K. E. PUTTICK (*British Iron and Steel Research Association*). A little while ago when growing single crystals and bicrystals of tin I made some observations, which though they were made during the course of another investigation and not followed up in any way, I think are worth mentioning here.

The single crystals were grown in Pyrex tubes and in moulds of graphite and sindanyo ; the latter were split cylinders about 1·5 in diameter with two holes for the metal, each about 0·8 cm diameter and 14 cm long. The crystals grown in Pyrex were of comparable size. The bicrystals were grown by Chalmers' method (1940) in boats of graphite or sindanyo. The speed of furnace travel was about 10 cm/h and the average temperature gradient, though not measured, was of the order of 10 deg/cm.

It was found that there were differences between crystals grown in these different materials. First, the lineage or macro-mosaic structure was found to be very marked in those crystals grown in glass, with respect both to the relatively wide range of angles in the crystallites and to their small size. In crystals grown in graphite and sindanyo, this substructure, though usually observable, appeared to have wider bands and smaller dispersion of orientation. Figure 1 shows a Laue photograph of a crystal grown in graphite which shows no evidence of macro-mosaic: in fig. 2, grown in sindanyo there is only slight splitting of spots. Figure 3 and 4 of crystals grown in glass, show that the crystallites were small relative to the diameter of the x-ray beam and there is a range of angles of order 10°. In fig. 3 the disorientation of the structure seems to correspond to rotation about a certain direction. This was frequently found to be the case, the crystallites forming a fibrous structure with the specimen direction as the common axis. Figure 4, however, shows a substructure which cannot be so described.

The orientation of different crystals was observed in all cases to have a most probable value for the *c*-axis of about 90° relative to the specimen axis, but the dispersion of orientations about this value was much greater for specimens grown in graphite than for those grown in glass. The orientation of the *a*-axis showed no tendency to a particular value. (These crystals showed no sign of polycrystals at the end of the rod as did some of those reported by Goss and Weintroub.)

In growing bicrystals from two single crystal seeds, symmetrically disposed with respect to the axis of the specimen, it was found that with graphite moulds the boundary, while being under these conditions roughly parallel to the specimen axis, was irregular. In sindanyo moulds, however, the boundary was microscopically straight. Straight boundaries confined to the median plane of the specimen were only consistently obtained when the *c*-axes of the seed crystals were perpendicular to the specimen axis.

It seems likely, then, that the thermal properties of the mould have a considerable influence on the growth of crystals, presumably because they affect the temperature gradient and rate of growth. Teghtsoonian and Chalmers (1951) have confirmed that the rate of growth does affect the size of the macro-mosaic units. It would be particularly interesting to know whether within the limits imposed by other considerations, the rate of cooling at nucleation does exercise a significant effect on the orientation of crystals, relative to the specimen axis. One can see that such an effect should exist, if we assume that the probability of a particular orientation is a function of a characteristic time of formation of a stable nucleus. Altering the rate of cooling implies varying the period of time spent in the neighbourhood of the freezing point, and so affecting the probability of formation of a nucleus of any particular orientation.

Suppose, for example, that the orientation α of a crystallographic axis relative to the specimen axis for a number of crystals has a normal probability distribution about a mean α_m, with a standard deviation σ. The probability $P(\alpha)$ of a particular orientation is then inversely proportional to the probable time of formation of a nucleus of that orientation, and we suppose that nucleation occurs when this is comparable with the time spent in the temperature range $(\theta, \theta+\delta\theta)$

$$\text{i.e. } P(\alpha) \propto \{1/(dt/d\theta)\}\delta\theta \quad \text{or} \quad \sigma^{1}(2\pi)^{-1/2}\exp[-(\alpha-\alpha_m)^2/2\sigma^2]=b(d\theta/dt)$$

where b is a constant. Putting $\alpha=\alpha_m$ we see that the standard deviation of orientation is inversely proportional to the rate of cooling at nucleation. The question we have to answer, then, is: why is there a minimum nucleation time characteristic of a certain orientation? One possible reason is that the surface energy is anisotropic, so that the freezing point and rate of growth are functions of orientation (Chalmers 1949).

This picture, however, offers no explanation of the curious fact that the conditions which apparently reduce the range of orientations between different crystals increase the range of the substructure orientation.

Chalmers, B., 1940, *Proc. Roy. Soc.* A, **175**, 100; 1949, *Ibid.*, **196**, 62.
Teghtsoonian, E., and Chalmers, B., 1951, *Canad. J. Phys.*, **29**, 370.

Asymmetry of Texture in Cubic Crystals Grown from the Melt

By P. R. ROWLAND
Department of Biochemistry and Chemistry, Guy's Hospital Medical School, London

MS. received 10*th March* 1952; *read before the Physical Society at Southampton on* 19*th December* 1951

ABSTRACT. Careful electro-polishing of single crystals pheres of copper (face-centred cubic system) grown from the melt has revealed that they appear to consist of a number of parallel crystallites of square cross section with their lengths parallel to the [001] axis which lies nearest in direction to the ' direction of growth ' in the melt.

Single crystal copper spheres about 15 mm in diameter on shanks (6·28 mm diameter) were carefully turned from cylindrical ingots which had been formed by solidification of the melt from the bottom upwards. They were then anodically polished to remove the superficial layer of polycrystalline material left by turning. This was done very carefully in a bath of 50% v/v H_3PO_4 and at current densities of about 1 amp/cm², holding the shank slantwise, rotating the crystal in the bath and dipping it in and out continually so that scoring due to the movement of bubbles upwards or the downward movement of dense electrolyte near the anode ('tear' formation) was minimized. This treatment aimed at producing as uniform a surface as possible. After about 15 minutes, when all polycrystalline material had been removed, dimples or pits were observed on the two cube faces which had been uppermost and lowermost respectively during growth of the crystal from the melt. The dimples lay along lines parallel to the cube face sides of the generating cube, forming a sort of cross-hatched pattern. On the other four cube faces, which lay nearer the vertical during growth of the crystal, the dimples lay along lines parallel to that cube edge direction which was nearest the vertical.

This is illustrated in figs. 1 and 2. The line AB in fig. 1 (*a*) is the direction of steepest temperature gradient and the line CD that of the nearest cube edge direction to it. Figure 1 (*b*) indicates the lines along which the dimples would lie on a cubic crystal cut from the ingot. Figure 2 shows the dimples on a spherical crystal. The line CD is the same in all three illustrations.

The dimples were about 0·1 mm diameter with about 0·3 mm between rows. The rows were not perfectly straight on the uppermost and lowermost faces but waved slightly rather like the markings in a finger print. Cracks could be observed on the sides of the crystal on the {011} (cube edge) regions. They were about 0·45 mm apart and appeared to form the boundaries of crystallites. In general, the dimples tended to avoid the cracks rather than to coincide with them. The only conclusion that can be drawn from these observations is that the crystals grow along the cube axis nearest to the direction of steepest temperature gradient, rather than along the temperature gradient itself. It will be interesting to note whether similar non-equivalence along main crystal axes is observed in the mechanical, magnetic and electrical properties of metal crystals.

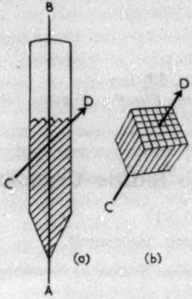


Fig. 1.

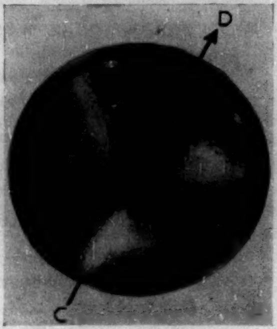


Fig. 2.

It is generally supposed that crystal growth occurs at dislocations and it is therefore to be expected that the direction of growth would be a singularity since these could hardly be randomly disposed. This may explain why the centres of faces in crystals growing from solution are frequently the points at which fresh layers commence (Bunn and Emmett 1949).

A very tentative explanation of the pattern is that the crystals grow as a series of closely aligned crystallites, each exposing cube facets and extending by deposition of material on the uppermost of these. This raises interesting speculations as to the initiation of the crystallites at their lower ends which must lie on the periphery of the ingot (see fig. 1 (*a*)). On the other hand, dendritic growth may give rise to the effect.

The author is indebted to Mr. A. J. Goss for drawing his attention to the fact that a similar structure, but on a rather finer scale, had been observed in the grains of polycrystalline copper-rich alloys by Northcott and Thomas (1939), using repeated mechanical polishing and etching. These authors assumed that the effect was due to dendritic growth but in that case it would have been expected that considerable intermeshing of the dendrites would have taken place at the boundaries between crystals. Most of their photographs show little indication of this.

Similar patterns have been observed by the author on the upper menisci of nearly all copper single crystals formed by upward growth from the melt. This is evidence that the effect is due to facet formation during growth, contrary to the result found by Anderson for zinc (Anderson 1932). Since our crystals were formed by very slow growth, the crystal–liquid interface may be expected to display the equilibrium form for the crystal, whereas Anderson allowed more rapid growth, and it has been shown by the author (unpublished) that this would tend to suppress facet formation.

The copper used was spectroscopic standard material* but, although grown in purified hydrogen, may have acquired a trace of oxygen, sufficient to give rise to an adsorption complex on the slowly growing crystal and hence to aid facet formation. Occasional occlusion of oxygen during this process may have given rise to the regions which become pitted on polishing. Alternatively, the pits may have centred on dislocations round which the growth of each crystallite took place.

The author wishes to thank Miss S. Treadgold and Mr. C. Engel for the illustration and photograph.

REFERENCES

Anderson, P. A., 1932, *Phys. Rev.*, **40**, 596.

Northcott, L., and Thomas, D. E., 1939, *J. Inst. Met.*, **55**, 205.

Bunn, C. W., and Emmett, H., 1949, *Discussions of the Faraday Society*, No 5 (Crystal Growth), p. 119.

The Growth of Thallium-Bromo-Iodide Crystals

By K. M. GREENLAND

British Scientific Instrument Research Association

MS. received 10*th March* 1952; *read before the Physical Society at Southampton on* 19*th December* 1951

ABSTRACT. Crystals of thallium-bromo-iodide, or KRS-5, are transparent to radiation of wavelengths between 0·5 μ and about 40 μ. This note briefly describes the growing of KRS-5 single crystals and their optical working. The crystals are grown in vertical cylindrical crucibles by a progressive freezing of the melt from bottom to top. The physical, chemical and mechanical properties of the material are summarized.

Thallium bromide and thallium iodide in the proportion 44:56 mole per cent form a mixed crystal, known as KRS-5, by solid solution (Koops 1948) which is transparent to radiations in the region 0·5 μ–40 μ. Since KRS-5 can be formed into optical components by grinding and polishing and is chemically stable, it is a very valuable addition to the range of optical materials.

Although KRS-5 is cubic in structure and therefore isotropic it is for various reasons desirable to have it in the form of single crystals; for instance, intercrystalline cracks are liable to form along the boundaries when the material is being ground or polished. The most suitable method of growing the single crystals appears to be that devised by Bridgman and applied to KRS-5 by

---

* Supplied by Messrs. Johnson, Matthey, Ltd., London.

Smakula and others (Elliott and Mayes 1945, Tuttle and Egli 1946). In this method the highly purified bromide and iodide, mixed in the proper proportions, is contained in a cylindrical crucible of platinum or Pyrex, of which the closed end is conical. The steady and progressive freezing through the melt which is necessary for the growth of a single crystal is encouraged by lowering the crucible, pointed end first, through two coaxial cylindrical furnaces. The temperature of the upper furnace is about 460° c, that is, some 45° c above the melting point of the mixture; the lower furnace is at roughly 330° c. The furnace temperature must of course be closely controlled to prevent a too rapid wandering of the region of freezing.

A steep temperature gradient in the region of freezing is an essential feature of the method. It has the obvious purpose of 'anchoring' the region of freezing and it is also believed to assist the rejection of impurities which might disturb uniform crystal growth or the optical properties of the crystal. In order to achieve a very steep temperature gradient we have tried lowering the crucible into a bath of liquid metal in the lower furnace. This also has the advantage that after solidification the crystal enters a region of nearly uniform temperature. However it seems that the temperature gradient near the freezing boundary may be too high and that the crystal structure, which is easily deformed at high temperature may be dislocated. At present, during modifications to the apparatus, we have returned to the use of empty air space in the lower furnace and the results are very promising. An annular baffle between the two furnaces helps to maintain a reasonably high temperature gradient. With this arrangement the rate of growth for a 2 in. diameter crucible is about 1 mm/hour.

When solidification is complete the crystal must be annealed and since it is plastic even at room temperature and is subject to delayed elastic effects the whole cooling process must be carried out very slowly.

KRS-5 has a density of 7·4 g cm^{-3} and a linear coefficient of expansion about seven times greater than that of glass. It has that peculiar combination of toughness and plasticity commonly described as 'horny'. It cannot be cleaved but it can be sawn, turned on a lathe, abraded with emery and optically polished with the usual polishing powders, though the technique of optical working differs in some ways from that for glass (Cooke 1947). In particular, flow lines formed below the surface during grinding appear as raised lines some time after the surface has been polished. To avoid this the blank is annealed again after grinding (Smakula and Klein 1950).

The refractive index of KRS-5 in the infra-red is 2·45–2·25; the dispersion falls steeply from 0·1 at $1\,\mu$ to 0·002 at $10\,\mu$, and above that wavelength it remains low (Plyler 1947, Tilton *et al.* 1949).

The solubility of KRS-5 in water is comparatively low (0·05 g/100 cm^3) and polished surfaces do not deteriorate in a humid atmosphere.

REFERENCES

Cooke, F., 1947, *J. Opt. Soc. Amer.*, **37**, 984 (A).

Elliott, A., and Mayes, E., 1945, *C.I.O.S. Report*, File 33-9, 8.

Koops, R., 1948, *Optik*, **3**, 298.

Plyler, E. K., 1947, *J. Chem. Phys.*, **15**, 885.

Smakula, A., and Klein, M. W., 1950, *J. Opt. Soc. Amer.*, **40**, 748.

Tilton, L. W., Plyler, E. K., and Stephens, R. E., 1949, *J. Res. Nat. Bur. Stand., Wash.*, RP.2008, **43**, 81.

Tuttle, O. F., and Egli, P. H., 1946, *J. Chem. Phys.*, **14**, 571.

Growth of Crystals of Caesium Bromide, and the Infra-Red Transmission Limit in Optical Materials

By A. C. G. MENZIES

Messrs. Hilger and Watts Ltd., Hilger Division, 98 St. Pancras Way, Camden Road, London N.W.1

MS. received 10*th March* 1952; *read before the Society at Southampton on* 19*th December* 1951

ABSTRACT. Following the development of methods for growing single crystals of materials having good optical properties, the search has been intensified for new substances. One such is caesium bromide, and this has been grown by the Kyropoulos–Pohl method of drawing from the melt; a modification of the standard process is described.

In the search for new materials, it is useful to have some way of predicting the useful long-wavelength transmission limit, and a rule is stated which forms a useful guide.

§ 1. INTRODUCTION

Crystals such as the alkali halides which are used as optical materials are for the most part grown by one of two methods. In one, which owes much to the work of Bridgman and of Stockbarger, the powdered material is moved slowly across a small finite temperature discontinuity maintained by two adjacent furnaces held at different temperatures.

In another process, developed by Kyropoulos and extended by Pohl and others, the crystal is drawn from the melt.

Some crystals are best grown by one of these two methods. In the writer's laboratory the second method has been used extensively for growing large crystals (9 in. diameter) of rocksalt and potassium bromide, and smaller ones of sodium iodide with added thallium, for scintillation counters. Occasionally small ones of sodium bromide, potassium iodide, silver chloride and rubidium bromide have been grown. In the growing of these materials a modification of the standard method has been developed, and this is described.

In the latter part of the paper is stated a rule which has been found useful for predicting whether the infra-red transmission of a material is likely to be favourable and therefore whether it is worth while to attempt to grow it.

§ 2. GROWING FROM THE MELT

A seed crystal is selected and mounted in a chuck, preferably of nickel or of nickel-plated material, which is water-cooled. The powdered material is placed in a crucible (often of unglazed porcelain) standing in a furnace. This furnace has a heating element at the bottom as well as the sides, to prevent the formation of a stalagmite. The seed crystal in its chuck is suspended over the top of the powdered material with the water running, and the heat is switched on. The material slowly melts, and for a 9 in. crystal this part of the process takes a week—hurrying leads to cracked crucibles and spoiled furnaces.

The seed is then lowered into the melt which is raised about 20° above the melting point, and the surface is melted off to form a new surface. The temperature is then lowered to a few degrees above the melting point and growth commences. A 'bun-shaped' mass forms on the seed, and extends sideways until it would seem to be in danger of reaching the wall of the crucible. At this stage, orthodox procedure is to raise the seed and its attached bun, and form a new bun on the underside. If this is done the finished crystal has a corrugated surface. It has been found in the writer's laboratory that by paying special attention to the heating arrangements, especially avoiding any asymmetrical cooling, and by proper selection of the size of the vessel, it is not necessary to lift the crystal, but one can rely on the liquid level dropping. This comes about because the solid is denser than the liquid. A smooth-faced crystal results, instead of the corrugated face normally obtained. This growing process also takes a week. Finally, another week is spent in the crystal cooling down.

While from first to last three weeks is necessary for a 9 in. crystal, a 3 in. one need only take about two days.

A special advantage of drawing from the melt is that one may observe the crystal from time to time to see if it is growing as a single or a poly-crystal, and, in the latter case, one may melt back and start again without great loss of time.

§ 3. CAESIUM BROMIDE

Powdered caesium bromide of high purity was obtained from Dr. Norman Wright, of the Dow Chemical Company, and was grown into a crystal by the method just described without any particular difficulty. A seed was made by melting some and allowing it to cool, and then picking out a sufficiently large piece from the cooled mass.

From the crystal was cut a sixty degree prism of good quality, and this is being used to measure the refractive indices of the material.

This material is likely to be important as a rival to thallium bromo-iodide as an infra-red material; it is likely to be usable to a wavelength in excess of 40 microns. Recent measurement by Plyler gives a transmittance of 76% at $38\,\mu$ for a 7 mm-thick window. Unlike the thallium compound, it is colourless and it has mechanical properties which, from the optical worker's point of view, are little worse than those possessed by potassium bromide. It is hygroscopic but not extremely so. The crystal was grown by Mr. J. Skinner.

§ 4. INFRA-RED TRANSMISSION LIMITS

The search for useful infra-red transmitting materials continues, and it is convenient to have some guide as to what one may expect beforehand. In compiling a list of wavelength limits of useful transmissions—by which is meant here the wavelength beyond which a spectroscopist would not seek normally to use a material—the author noticed a rule which so far appears to be obeyed by these materials. The upper wavelength limit of useful transmission is about one-third of the residual ray (reststrahlen) wavelength. This limit corresponds well with the wavelength at which the transmission of a two centimetre thick sample has dropped to about 10%.

The extent to which the rule is followed may be gauged from the following table :

Transmission Limits and Residual Rays

Material	Useful wavelength limit (μ)	Residual rays
Fused silica	3·6	9
LiF	6	17
CaF$_2$	9·5	32
NaCl	17	52
KCl	21	63
AgCl	23·5	82
KBr	27	82
KI	32	94
TlBrI	38	120
[CsBr	>38	<134]

The residual ray wavelength fits better than the infra-red absorption maximum, as may be seen, for example, by considering lithium fluoride. The residual ray wavelength for caesium bromide is so far not known, but the absorption maximum is 134 μ, and the residual ray maximum will be less than this, by an amount due to the Försterling correction. One may conclude that caesium bromide should have a useful limit in the above sense of just over 40 μ, and the likelihood of this being so seems to be indicated by the measurement of Plyler referred to above.

Crystal Growth Ideas and Crystallographers*

The writer was brought up to include, in the latter term of the title, anyone who worked on the enormous variety of problems connected with the properties of crystals ; optics, hardness, glide and slip, chemical failure, solubility of discrete crystal planes, crystal structures as evaluated by x-rays and similar methods, and the many problems and difficulties associated with their growth. It has not escaped the attention of the writer that a ‘ crystallographer ’, at least in England, need be conversant with only one of these categories and can be quite ignorant of anything else happening around him, indeed, need not be very conversant with the elements of descriptive morphology known for over a hundred years. How else would it be necessary to refute the *current* implications of a theory, possibly valid without these latest frills, wherein a phenomenon occurring at times on a face which scarcely grows, is supposed to supply the ‘ motif ’ of growth for all the crystal. The side facets, on which most growth occurs, are dismissed as ‘ kinked ’ or ‘ incomplete ’ faces, though $\{10\bar{1}0\}$ of either CdI$_2$ or SiC is quite as complete as $\{0001\}$, it being forgotten that a *unit-cell*, e.g. of a hexagonal crystal, must be clothed in $\{10\bar{1}0\}$ as well as $\{0001\}$. What we require, then, is a theory which makes them grow more rapidly, and solid phase to separate via them, as against those on which spirals appear, and on which, from the very shape of the completed crystal, anyone versed in elementary morphology would immediately recognize that little or no addition to the crystal had taken place during growth.

Briefly, in addition to the beautiful observations of Verma (1951) and Amelincx (1951) on SiC and further ones on CdI$_2$ by Forty (1951), on mica by Amelincx (1952), on colloidal gold by Forty (1952) and AlB$_2$ from Al (Horn *et al.* 1952), the writer has observed all the phenomena on a scale up to nearly a millimetre between loops, both closed and straight-edged, all practically visible to the naked eye but not commented upon by others. He appeals to anyone to examine for himself a lump of SiC crystals—not specially selected ones—and verify that the incidence of reasonably good crystals is (in our present sense) far above one in ten : that really good specimens are far from common and that the great bulk of SiC crystals are dendritic, often curved-surface masses, and it would be difficult, with such unreliable material as this, to connect up with any theory of crystal growth. The writer has all along considered that some vortex of vapour, possibly containing the material for growth, was present at the conclusion of growth. This formed a smear of SiC at the end and imposed some kind of spiral on it. Often (fig. 1), the spiral-covered dendrite is visible. The crystal is punctured in depth by holes of various shapes and these are covered with a thin smear of SiC, the latter being detected from more solid SiC underneath through total reflection at the surface SiC–air, rendering the crystal paler in reflected light at these numerous places. The deposition of the spiral is, in this view, an event occurring near the end of crystallization.

---

* Read at Physical Society Meeting at Southampton on 19th December 1951.

The picture of CdI_2 (fig. 2) shows excessively thin sheets, 500 Å thick by ½ mm across and several thousands of times wider than they are thick. Negligible amounts enter the crystal by growth on such {0001} surfaces. Even in an average CdI_2 crystal the sides are practically responsible for over 99% of the solid phase. While the drop evaporated (3 hours) the shade of Newton's colours (straw, 1st, and green, 2nd order, latter bigger) persisted, but the triangular seed observable on the straw-coloured one fringed itself in, all this time, with red, 1st order, and became itself blue, 2nd, the crystal growth, all this time, on the bulk of {0001} was nil. In practically all cases given so far, we know nothing of how the crystal grows; alternatively, we artificially restrict it in some way (coverslip with CdI_2, innumerable cavities and vapour streams in SiC). Mica grows along its optimum direction which is along a sheet from side to side. How can a *growth* spiral come about under such circumstances? There is something suspicious in all other cases quoted so that bona-fide crystal growers drop the idea like hot cakes as inapplicable to their work. In the writer's opinion the early views of Frank (1949) have something to be said for them. The surface spirals and other similar features are also well attested as a phenomenon, so far unexplained, but to make the former theory depend upon the latter observations does not appear to the writer to make sense.

Figure 1. A straight-forward spiral on SiC, visible with a hand-lens.

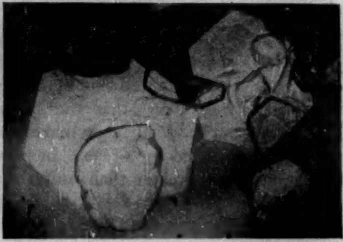


Figure 2. Excessively thin plates of CdI_2 showing Newton's lowest orders of colour, formed under a cover-slip.

Crystallographic Department,
University of Manchester.
10th March 1952.

H. E. Buckley.

Amelincx, S., 1951, *Nature, Lond.*, **167**, 940; 1952, *Ibid.*, **169**, 580.
Forty, A. J., 1951, *Phil. Mag.*, **42**, 672; 1952, *Ibid.*, **43**, 502.
Frank, F. C., 1949, *Discussions of the Faraday Society*, **5**, 49.
Horn, F. H., Fuller, E. F., and Kasper, J. S., 1952, *Nature, Lond.*, **169**, 927.
Verma, A. R., 1951, *Nature, Lond.*, **167**, 939.

Other papers read at this meeting have been published elsewhere as follows:

Lawson, W. D., 1951, 'A Method of Growing Single Crystals of Lead Telluride and Lead Selenide,' *J. Appl. Phys.*, **22**, 1444; 1952, 'Oxygen-free Single Crystals of Lead Telluride, Selenide and Sulphide,' *J. Appl. Phys.*, **23**, 495.

Optics of the Focusing X-Ray Microscope

By J. DYSON

Research Laboratory, Associated Electrical Industries Ltd., Aldermaston Court,
Aldermaston, Berks.

Communicated by T. E. Allibone; *MS. received* 10*th December* 1951, *and in amended form* 8*th April* 1952; *read before the Society* 2*nd May* 1952

ABSTRACT. A brief description is given of Kirkpatrick and Baez' x-ray reflecting microscope. The types of aberration possible in optical systems without axial symmetry are discussed and the effect of spherical aberration on the diffraction pattern of an object point is evaluated. The resolving power of a circular mirror is calculated, and a method of correcting field obliquity is described. A method is given for the design of a figured mirror to give smaller spherical aberration with zero field obliquity.

Design equations are given for a system employing two circular mirrors in each plane, by which spherical aberration, obliquity and curvature of field can be corrected. The computed performance of the system is described.

§ 1. INTRODUCTION

It has been pointed out (Jentzsch 1929, Kirkpatrick and Baez 1948, Kirkpatrick 1950) that the phenomenon of total external reflection of x-rays by solid surfaces makes possible the formation of real images of objects illuminated by x-rays. In particular, such images may be formed with a magnification greater than unity, thus suggesting the possibility of an x-ray microscope.

The scheme proposed by Kirkpatrick and Baez is illustrated in fig. 1. A small object illuminated by x-rays is placed near a concave mirror so disposed that the central ray of the x-ray beam strikes the mirror at a glancing angle i. If the angle be sufficiently small, total reflection of the x-rays takes place at the surface of the mirror, which exerts a focusing action on the beam. The image is formed at a distance from the mirror which is greater than that of the object, and under these conditions is larger than the object. However, the image is strongly astigmatic, and to overcome this defect a second reflector is placed in the x-ray beam with its plane perpendicular to that of the first, and so arranged that the astigmatism of the first mirror is cancelled by that of the second. The magnification of the image is then different for two mutually perpendicular directions. Kirkpatrick and Baez outline a method for overcoming this defect, but in any case it is not serious and can be corrected later in the optical enlarger.

Apart from any geometrical aberrations which may arise, the attainable resolution is given by the maximum aperture angle and the x-ray wavelength. For large magnifications the emergent rays will be nearly parallel, and under these conditions it is easily seen that the maximum semi-aperture angle is equal to the glancing angle i. The resolving limit is therefore equal to $\lambda/2i$ where λ is the x-ray wavelength.

The maximum glancing angle for which total reflection will take place from a given substance is, in first approximation, proportional to the wavelength, so that the resolving limit tends to be independent of the wavelength for a given mirror material. Kirkpatrick and Baez give $i = 7 \times 10^5 \lambda$ (where λ is in cm) for the glancing angle from the platinum metals which, they state, give the largest angles for a given wavelength of any known substance. This value gives a minimum resolving limit of 70 Å, but this cannot be approached closely in practice because of the geometrical aberrations of the system. Prince (1950) has calculated the resolution on the basis of the dimensions of the geometrical confusion patch for a system using mirrors of circular longitudinal cross section, and estimates the resolving limit to be in the region of 3 400 Å.

The effects of surface conditions and accuracy of the mirrors have been discussed by Jentzsch (1929), Ehrenberg (1947, 1949 a, b) and Lucht and Harker (1951).

The object of this paper is to examine the types of aberrations which can arise in a system of mirrors used at very small glancing angles, and to indicate means by which such aberrations may be corrected or reduced.

§ 2. ABERRATIONS OF NON-AXIAL SYSTEMS

The most noticeable feature of optical systems such as that illustrated in fig. 1 is that there is no axis of symmetry as in the case of more conventional systems. This suggests that the possible types of aberrations may be different from those occurring in more familiar cases.

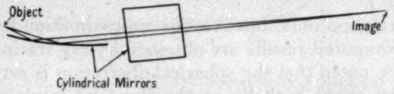


Fig. 1. Schematic arrangement of the Kirkpatrick–Baez microscope.

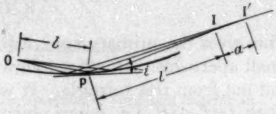


Fig. 2. Spherical aberration.

Considering one plane at a time, the position of emergence of a ray from the exit pupil can be described by a single coordinate x. Rays with the same value of x but separated in a direction perpendicular to the plane of incidence on the mirror of the central ray will suffer the same aberration if the mirror be of cylindrical shape, and, if it be spherical, the aberration will differ only by an insignificant amount for the values of angular aperture considered here. Also, as the deflections produced by the two mirrors are at right angles to each other, interactions between them are negligible for small aperture and fields. The distance of the object point from some arbitrarily chosen origin, measured at right angles to the imaging beam (which is assumed small in angular aperture) can be designated by a 'field coordinate' h. Then the angular aberration of the ray can be conveniently expressed as a series of products of powers of x and h:

$$\text{Angular aberration} = a_{10}x + a_{20}x^2 + a_{30}x^3 \;\ldots\; + a_{01}h + a_{11}xh + a_{21}x^2h + a_{31}x^3h + \;\ldots\; + a_{02}h^2 + a_{12}xh^2 + a_{22}x^2h^2 + a_{32}x^3h^2 + \;\ldots \qquad \ldots\ldots(1)$$

In a system with axial symmetry some of the terms in (1) would be prohibited by symmetry considerations, but here all the terms may have finite values. The first few terms may be interpreted as follows:

$a_{10}x$, change of focal plane; $a_{20}x^2$, primary spherical aberration; $a_{30}x^3$, secondary spherical aberration; $a_{01}h$, change of magnification; $a_{11}xh$, obliquity of field; $a_{21}x^2h$, coma; $a_{02}h^2$, distortion; $a_{12}xh^2$, curvature of field.

The term 'spherical aberration' is used for $a_{20}x^2$, $a_{30}x^3$, as these quantities depend only on the aperture. The primary spherical aberration varies with the aperture in a manner similar to that of the tangential component of coma in axially symmetrical systems, but does not vary with the field.

§ 3. SPHERICAL ABERRATION

Consider a mirror of circular cross section and radius of curvature R on which is incident an infinitely narrow pencil of rays diverging from an object point O (fig. 2), distant l from the point of contact P of the principal ray with the mirror. The pencil is focused at a point I, distant l' from P. Then, if the glancing angle made by the principal ray with the mirror be i, the distances l and l' are connected by the relationship $1/l+1/l'=2/iR$. A ray emanating from O and striking the mirror at a small but finite distance r from P will intersect the principal ray at a point I′, distant a from I. Kirkpatrick and Baez give an expression for the distance a which may be put into the rather more convenient form

$$a=-\frac{3rl(l-iR)}{(2l-iR)^2}=-\frac{3rl'(iR-l')}{i^2R^2} \qquad \ldots\ldots(2)$$

or, in terms of the aperture angle α between the aberrant ray and the principal ray,

$$a=-\frac{3\alpha l'l(l-iR)}{i(2l-iR)^2}=-\frac{3\alpha l'^2(iR-l')}{i^3R^2}. \qquad \ldots\ldots(3)$$

This gives the primary spherical aberration only, but is a fair approximation for small apertures. In what follows, computed results are obtained by ray tracing and not from this formula. It will be noted that the spherical aberration is zero when $l=l'=iR$, i.e. when object and image both lie on the Rowland circle of the mirror (the circle of radius $R/2$, passing through the centre of curvature of the mirror and tangent to it at P). This is, then, a particularly favourable case for image formation, but as the magnification is unity it is of little interest in microscopy.

The size of the geometrical aberration patch is equal to αa, but in the practical case the distribution of intensity is modified by diffraction effects, and it is necessary to take these into account.

If the longitudinal aberration is equal to $k\alpha$ the deviation from circularity of the wave front for any angle α is equal to $k\alpha^3/3$.

The intensity at a point in the image plane through I and normal to PI is given by

$$I=\frac{1}{4\alpha^2}\left[\int_{-\alpha}^{\alpha}\cos\frac{2\pi}{\lambda}\,(k\alpha^3/3+S\alpha)\,d\alpha\right]^2, \qquad \ldots\ldots(4)$$

where S is the distance of the point considered from I, and the unit of intensity is the value at the focus if the system were free from aberration. For $k=0$ (4) gives the well-known result $I=\sin^2(2\pi S\alpha/\lambda)/(2\pi S\alpha/\lambda)^2$. The integral has been evaluated numerically for values of $k\alpha^3/3$ equal to $\lambda/4$, $\lambda/2$ and $3\lambda/4$. The results are shown in fig. 3.

The most obvious feature of the curves as the aberration increases is a progressive shift of the maximum towards the side to which the rays are deviated, as is to be expected. There is also a fall of the maximum value of the intensity, a progressively smaller fraction of the total energy being concentrated in the central lobe and a corresponding increase of intensity in the side lobes on the side to which the central lobe is displaced. The lobes on the other side are diminished in height, but a rather unexpected phenomenon is the development of a long plateau or 'toe' to the intensity curve on this side.

Since the width of the central lobe does not vary greatly for moderate amounts of aberration, it is evident that the resolution is not profoundly affected. The contrast will diminish, however, by reason of the large side lobes, and this effect may well be of importance if the instrument is to be used for the measurement of differential x-ray absorption, as, for example, in the location of specific elements (Engstrom 1947, 1949). In this case it is evidently desirable to keep the spherical aberration down to the lowest possible value.

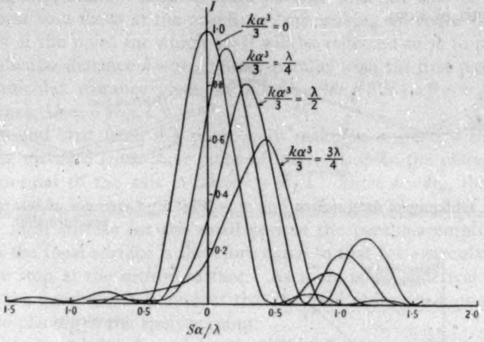


Fig. 3. Intensity distribution in focal plane for varying amounts of spherical aberration.

For a mirror of circular cross section, and of focal length f, used with a glancing angle of i for the principal ray, the wave front distortion p for a semi-aperture angle α and infinite magnification is given by $p=f\alpha^3/4i$. If a quarter of a wavelength of spherical aberration be allowed, the resolving limit d is given by

$$d=\tfrac{1}{2}(f\lambda^2/i)^{1/3}. \qquad \ldots\ldots(5)$$

To obtain the variation with wavelength, the aperture angle is obtained by substituting $i=7\times10^5\lambda$ in (5), when the resolving limit is given by $d=0{\cdot}0056(\lambda f)^{1/3}$. For $i=0{\cdot}01$, $\lambda=1{\cdot}4$ Å and $f=1$ cm, d is equal to 1 350 Å. It is evident from (5) that this value changes only slowly with the wavelength or the scale of the system.

§ 4. OBLIQUITY OF FIELD

Kirkpatrick and Baez point out that a circular mirror used in this manner gives a focal surface which is strongly inclined to the normal to the principal ray. However, because of the large spherical aberration, the inclination of the focal surface for ray pencils of small aperture is a function of the position of the aperture stop. For the particular case of infinite image distance, as shown in fig. 4, the angle β between the focal surface and the principal ray is given for small apertures by

$$\tan\beta=i(iR-2D)/(iR+D), \qquad \ldots\ldots(6)$$

where D is the distance, measured in the space of the short conjugate, between the point of incidence on the mirror and the aperture stop.

In the case considered by Kirkpatrick and Baez, the aperture stop is the mirror itself, i.e. $D=0$; the focal surface thus makes an angle i with the principal ray. This is evidently a very unfavourable condition, for the object would have to lie in a plane almost coinciding with the line of sight. However, it can be seen from (6) that β becomes 90° if $D=-iR$. This corresponds to a virtual stop on the short-conjugate side, but the same effect is produced by a real stop on the long-conjugate side at a distance $iR/3$ from the mirror, which is the image formed by the mirror of the required virtual position. The system is then as shown in fig. 5.

The focal surface will, of course, only be flat in first approximation, and computation shows that the sharp field is limited by field curvature. Figure 6 gives the results of computation on a system for which $R=200$, $i=0{\cdot}01$, giving a focal length of 1 cm.

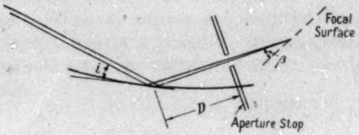


Fig. 4. Obliquity of focal surface.

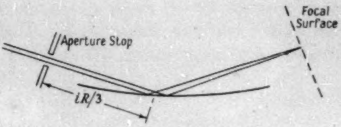


Fig. 5. Correction of field obliquity by aperture stop.

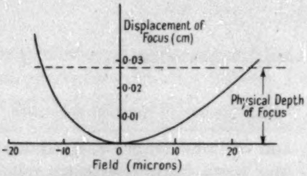


Fig. 6. Field curvature of circular mirror with aperture stop.

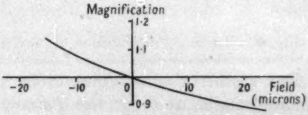


Fig. 7. Variation of magnification with field.

As the glancing angle varies over the field, the system must be used at a wavelength somewhat greater than the 1·4 Å suitable for $i=0{\cdot}01$. Taking $\lambda=1{\cdot}6$ Å, which is suitable for the field shown, the semi-aperture angle α for a quarter wavelength of aberration is $0{\cdot}545\times10^{-3}$, giving a resolution $d=1\,500$ Å. The physical depth of focus $\lambda/2\alpha^2$ is then 0·0275 cm. It is seen that the system is suitable for a field of 37·5 microns. The mirror is about 0·26 cm in length, and the maximum length contributing to any one image point is 0·062 cm.

This may be compared with the case of a circular mirror with the aperture stop at the mirror surface. Taking $i=0{\cdot}01$, the physical depth of focus is reached at a total field of 5·5 microns, i.e. the sharp field is very small in extent. Reasonably extended fields can therefore only be obtained by employing much smaller aperture angles.

The aperture angle varies somewhat over the field for a fixed size of stop, so the depth of focus also varies; however, its variation is such as to give a somewhat larger field than the 37·5 microns quoted above. It follows that the magnification also varies over the field, but this is not a serious defect. Its variation is shown in fig. 7, where the magnification is given in terms of its value at the point where the glancing angle has the value 0·01.

§ 5. FIGURED MIRRORS

It is evident that spherical aberration can be removed completely if a parabolic mirror be used in the case of infinite image distance, but the field obliquity is very large. The radius vector from the focus of the parabola must meet the curve at a very small glancing angle, i.e. the latus rectum of the parabola must be very small compared with the radius vector of the portion of the curve used. Under these conditions the parabola can be represented with sufficient accuracy over the portion used by the polar equation $r=4r_0i^2/(2i+\alpha)^2$, where i is the glancing angle of the principal ray from the focus and α is the semi-aperture angle taken as a running coordinate. Incident rays parallel with the axis of the parabola will then come to a focus at the origin. A ray making an angle θ with the axis and incident at the point for which $\alpha=0$ will be reflected so as to pass the focus at a perpendicular distance $h=\theta r_0$. A ray parallel with the first passes the focus at a perpendicular distance given by $h+\Delta h=4\theta r_0 i^2/(2i+\alpha)^2=\theta r_0(1-\alpha/i)$ very nearly. Hence $\Delta h=-\theta r_0\alpha/i$.

This ray and that from the point $\alpha=0$ make an angle α with each other; therefore the distance from their point of intersection to the plane through the focus and normal to the axis is $\Delta h/\alpha=-\theta r_0/i$. Since $k=\theta r_0$, the intersection lies on a plane passing through the focus and making an angle $\beta=i$ with the axis. This is the focal surface for the small part of the parabola employed, and the obliquity of the focal surface is therefore equal to that for a circular mirror with the aperture stop at the mirror surface. As there is no spherical aberration in the case of the parabola, it is evident that the field obliquity cannot be rectified by a suitable placing of the aperture stop.

This suggests that by departing somewhat from the parabolic form it might be possible to design a combination of mirror and stop which, while possessing smaller spherical aberration than the circular mirror, could be given an acceptable value of field obliquity.

The method by which this can be achieved is as follows (fig. 8). The surface AB is represented with reference to the point F by the polar equation

$$\Psi=2ir^{-1/2}-f(r), \qquad \text{......(7)}$$

where Ψ is the angle between the radius vector and a fixed direction, and $f(r)$ is small compared with Ψ. The line FP is taken as a principal ray and FP assumed equal to unity. The 'unperturbed' curve $\Psi=2ir^{-1/2}$ then represents a parabola. The glancing angle i' between the radius vector and the curve is given by

$$i'\simeq\tan i'=-r\frac{d\Psi}{dr}=\frac{i}{r^{1/2}}+rf'(r). \qquad \text{......(8)}$$

The angle θ made with the axis by a ray from F after reflection is

$$\theta=\Psi-2i'=-f(r)-2rf'(r).$$

The aperture stop is taken to be at S, distant l from F, and the focal surface is taken to be a plane CD through F and normal to FP.

The principal ray for an object point F′, distant k from F, is SF′, cutting the curve in P′. Then $\mathrm{F'P}\simeq\mathrm{FP'}=r'$, say. The angle made with the axis by this ray after reflection is

$$\theta'=-f(r')-2r'f'(r')+\mathrm{F\hat{P}'F'}=-f(r')-2r'f'(r')+h/r'. \qquad \text{......(9)}$$

.B
5
2

Another ray, F′P″ from F′, meeting the curve at F″, where $F'P'' \simeq r' + \Delta r'$, is, after reflection, inclined to the axis at an angle

$$\theta' + \Delta\theta' = -f(r' + \Delta r') - 2(r' + \Delta r')f'(r' + \Delta r') + \frac{h}{r' + \Delta r'}.$$

Then the angle between the rays F′P′ and F′P″, after reflection, is

$$\Delta\theta' = -f(r' + \Delta r') - 2(r' + \Delta r')f'(r' + \Delta r') + \frac{h}{r' + \Delta r'} - \left[-f(r') - 2r'f'(r') + \frac{h}{r'}\right]$$

$$\simeq \Delta r'\left[3f'(r') + 2r'f''(r') + \frac{h}{r'^2}\right]. \qquad \ldots\ldots(10)$$

The degree of approximation is such that (10) takes account of variations of $\Delta\theta'$ which are proportional to $\Delta r'$ but neglects those due to the square and higher powers of $\Delta r'$, i.e. (10) gives the extent to which the focal surface conjugate to a plane at infinity deviates from CD but does not take account of spherical aberration. From (10) the condition for CD to be the true focal surface for parallel rays is

$$3f'(r) + 2rf''(r) + h/r^2 = 0, \qquad \ldots\ldots(11)$$

omitting the primes.

The solution of (11) is complicated by the fact that the relationship between h and r involves $f(r)$. The resulting differential equation can be solved, but leads to a power series which is only very slowly convergent and is of little practical value. However, an approximate solution may be found by numerical means. The steps are as follows: assuming the unperturbed curve $\Psi = 2ir^{-1/2}$, calculate h for a suitable range of values of r and an assumed value of l; calculate h/r^2 for each r and approximate by a polynomial of form $h/r^2 = \Sigma A_n r^n$; substitution in (11) and solution of the resulting differential equation gives $f(r)$ directly; substitute $f(r)$ in (7) and check the design by computation. This has been done for $l = 1$ and $i = 0{\cdot}01$. The resulting equation for the curve is

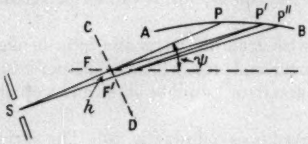


Fig. 8. Figured mirror.

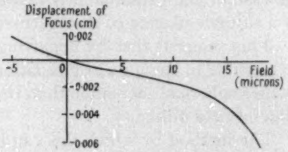


Fig. 9. Field distortion of figured mirror.

$$\Psi = 0{\cdot}02r^{-1/2} - 0{\cdot}0012457(\tfrac{1}{3}r - 1/15r^3). \qquad \ldots\ldots(12)$$

The computed form of field is shown in fig. 9. The spherical aberration is nearly constant over the field and is such that a semi-aperture angle of $0{\cdot}77 \times 10^{-3}$ can be used with a wavelength of 1·4 Å, giving a resolving limit of 900 Å, and a corresponding depth of focus of 0·0118 cm; thus the field is adequately flat over at least the extent shown. The computation was carried out using a two-term approximation to h/r^2. It is evident that further refinement of the design would give a better approximation to flatness of field. The total length of mirror employed for a focal length of 1 cm and a field of the extent shown in fig. 9 is approximately 0·7 cm.

It is found that the chief limitation on the design is that in order to reduce the spherical aberration the aperture stop must be nearer the focal surface, and hence a greater length of the mirror surface must be used for a given field. This is undesirable, for not only must a greater area of mirror be finished accurately, but also the ratio in scale of the two mirrors in mutually perpendicular planes must be greater, with a corresponding increase in the spherical aberration of the large mirror.

§ 6. SYSTEMS EMPLOYING TWO MIRRORS IN EACH PLANE

Harker (private communication) has pointed out that the field obliquity can be corrected by using two mirrors in each plane, with the concavities facing in opposite directions. An analysis of such compound systems shows that if the concavities face in the same direction both field obliquity and primary spherical aberration can be corrected. The mirror arrangement is shown in fig. 10.

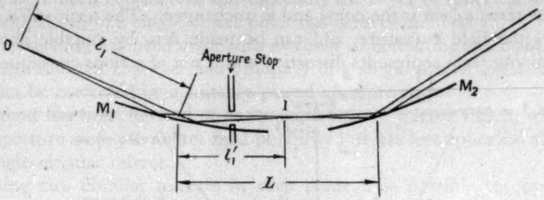


Fig. 10. Two-mirror system.

The object point O is imaged by the mirror M_1 at I and the second mirror M_2 forms an image of I at infinity. The radii and glancing angles of the principal ray are R_1, R_2 and i_1, i_2 for the mirrors M_1 and M_2. To simplify the algebra the aperture stop is taken to coincide with M_1. Under these conditions the angle between the principal ray and the focal surface for M_1 can be shown to be $\tan\beta = R_1 i_1^2/2l'$. For the second mirror eqn. (6) can be applied, setting D equal to L, the distance between the points of incidence of the principal ray. Then, for the overall field obliquity to be corrected

$$\frac{R_1 i_1^2}{2l'} = \frac{i_2(2L - i_2 R_2)}{L + i_2 R_2} = \frac{2 i_2 l_1'}{l_1' + 3 i_2 R_2/2}, \qquad \ldots\ldots(13)$$

where L has been replaced by $l_1' + i_2 R_2/2$, the condition for infinite final image distance. Equating the spherical aberrations of the two mirrors,

$$R_2 = 8 l_1'^2 (i_1 R_1 - l_1')/R_1^2 i_1^3.$$

Substitution of this value in (13) gives $i_2 = R_1^2 i_1^3 / 4 l_1' (3 l_1' - 2 R_1 i_1)$. Thus, as an example, if $i_1 = 0{\cdot}01$, $R_1 = 100$ cm, $l_1' = 0{\cdot}75$ cm, we find $R_2 = 112{\cdot}5$ cm, $i_2 = 0{\cdot}0133$, $L = 1{\cdot}5$ cm, $l = 1{\cdot}5$ cm.

The performance of this system has been ascertained by computation. The residual angular spherical aberration is very small, amounting to about $1{\cdot}7 \times 10^{-6}$ for a semi-aperture angle of 0·001 for a point in the centre of the field, and is approximately proportional to the fourth power of the aperture. Little use, however, can be made of this very small value of spherical aberration because of the field aberrations. The most important of these are coma and curvature of field.

The field curvature can be corrected with the help of the coma if the aperture stop be placed at a position different from that of the first mirror. To explain this it should be noted that from eqn. (1) it is possible to write the combination of the two aberrations as

$$\text{Angular aberration} = a_{21}x^2h + a_{12}xh^2. \qquad \ldots\ldots(14)$$

If the aperture stop be moved from M_1 but kept centred on the principal ray from the object point at zero field the beam from this point is not affected, but the principal ray from any other point with field coordinate h is now moved laterally in the beam emanating from that point by an amount proportional to h. This can be taken into account in (14) by replacing x by $X = x - kh$, where k depends on the new position of the stop. Equation (14) then becomes

$$\begin{aligned}\text{Angular aberration} &= a_{21}(x-kh)^2h + a_{12}(x-kh)h^2 \\ &= a_{21}x^2h + xh^2(a_{12} - 2a_{21}k) + h^3(a_{21}k^2 - a_{12}k). \qquad \ldots\ldots(15)\end{aligned}$$

In (15) the term $a_{21}x^2h$ is the coma and is unchanged. The term $xh^2(a_{12} - 2a_{21}k)$ represents the field curvature, and can be made zero by suitable choice of k. The remaining term represents distortion and is not of serious consequence.

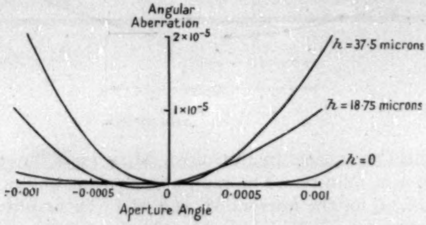


Fig. 11. Aberrations of two-mirror system.

The most suitable position for the stop is between the mirrors and distant approximately 0·4 cm from the point of incidence of the principal ray on M_1.

Under these conditions the angular aberration is plotted in fig. 11 for three values of h.

The maximum value of the glancing angle is 0·0143, corresponding to a wavelength of 2·06 Å. With this value, and allowing spherical aberration $\lambda/4$ at a field of 37·5 microns, the useful semi-aperture angle is $0{\cdot}89 \times 10^{-3}$ radians, corresponding to a resolving limit of 1 150 Å. However, most probably it will be sufficient to allow the spherical aberration to reach $\lambda/4$ at the half-field of 18·75 microns; this makes the semi-aperture angle $1{\cdot}3 \times 10^{-3}$ and the resolving limit 800 Å.

§ 7. GENERAL DISCUSSION OF RESULTS

It can clearly be seen from the above that the resolving power of the x-ray microscope cannot easily be varied by more than a factor of about two from the value given by a simple pair of crossed circular mirrors. This was forecast by Prince, although his value for the resolving limit, which was based on geometric optics alone, appears to be a little pessimistic in view of the true distribution of light in the image as given by wave theory.

The chief advantage to be gained from the use of systems more complex than that of Kirkpatrick and Baez would then seem to lie in the possibility of correction of field obliquity. The single circular mirror with a correctly placed aperture stop may prove to be the most useful on account of its simplicity. It has the advantage that, as the number of reflecting surfaces is a minimum, scattering will also be a minimum; furthermore, it seems likely that circular (or spherical) surfaces can be more accurately made than figured ones. Very careful experimentation would be necessary to prove the reality of the difference between the 1 500 Å obtainable by this system and the lower values obtainable by more complex systems, though it may be that the additional factor of almost two to be gained is worth the effort.

§8. CONCLUSIONS

The aberrations characteristic of mirror systems used far from their axes of symmetry have been considered, and especially the effect of spherical aberration on the Airy pattern of a single object point has been evaluated.

An investigation has been made into methods of correcting these aberrations. It has been indicated that the field obliquity of a single pair of crossed circular mirrors can be corrected by a suitably placed aperture stop.

A method has been devised of designing a figured mirror which, when used with an aperture stop, eliminates field obliquity but has less spherical aberration than a single circular mirror.

By using two circular mirrors in each plane it is possible to correct field obliquity, field curvature and spherical aberration, and the design equations and computed results are given for such systems.

The more complex systems considered do not increase the resolving power by as much as a factor of two, however, and it seems possible that the most rewarding system to use will be the simple circular mirror with suitably placed aperture stop.

ACKNOWLEDGMENT

The author wishes to thank Dr. T. E. Allibone for permission to publish this paper.

REFERENCES

EHRENBERG, W., 1947, *Nature, Lond.*, **160,** 330; 1949 a, *J. Opt. Soc. Amer.*, **39,** 741; 1949 b, *Ibid.*, **39,** 746.

ENGSTROM, A., 1947, *Transactions of Instruments and Measurements Conference, Stockholm*, p. 71; 1949, *Acta Radiol., Stockholm*, **31,** 503.

JENTZSCH, F., 1929, *Phys. Z.*, **30,** 268.

KIRKPATRICK, P., 1950, *Nature, Lond.*, **166,** 251.

KIRKPATRICK, P., and BAEZ, A. V., 1948, *J. Opt. Soc. Amer.*, **38,** 716.

LUCHT, C. M., and HARKER, D., 1951, *Rev. Sci. Instrum.*, **22,** 392.

PRINCE, E., 1950, *J. Appl. Phys.*, **21,** 698.

The Focal Properties and Spherical Aberration Constants of Aperture Electron Lenses

By M. M. MacNAUGHTON*
Research Laboratory, Associated Electrical Industries Ltd., Aldermaston, Berks.

MS. received 24th September 1951, *and in amended form 13th February* 1952

ABSTRACT. The focal properties and spherical aberration constants of a number of aperture electron lenses of varying diaphragm thickness and varying strength were determined numerically using potential distributions obtained (*a*) from a rigorous solution due to Ollendorff, and (*b*) on a resistor network analogue. The focal lengths of 'weak' aperture lenses of finite diaphragm thickness agree closely with those given by Davisson and Calbick's formula for the 'infinitely' thin aperture lens; 'strong' aperture lenses of all thicknesses show systematic deviations from this formula. The primary spherical aberration constants of diverging aperture lenses are small, and of such sign that they cannot be used for the correction of the spherical aberration of converging electron lenses.

---

§ 1. INTRODUCTION

An aperture or 'pinhole' lens exists where a diaphragm pierced by a circular aperture separates two spaces with different electric field strengths E_1 and E_2. The geometry of such a lens is shown in fig. 1. When the first field strength $E_1=(U_0-U_1)/d_1$ between electrode A_1 and diaphragm D, of thickness T and aperture radius R, is greater than the second $E_2=(U_2-U_0)/d_2$, between diaphgram D and electrode A_2, the lens is divergent and, conversely, when the first field strength is less than the second, the lens is convergent. The aperture lens of the first kind is the only known practical electron lens which is divergent; this type of lens is found in most accelerating systems, such as those used in electron guns.

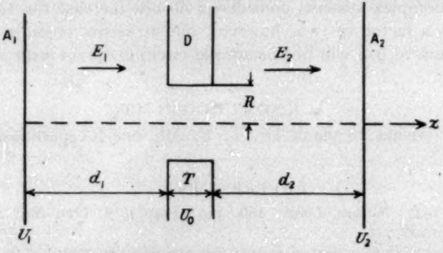


Fig. 1.

Davisson and Calbick (1931, 1932) have shown that an aperture lens, with an indefinitely thin diaphragm and with small field strengths E_1 and E_2, and a potential $U_0=U$ applied to the diaphragm, has the properties of a thin lens located in the plane of the diaphragm whose focal length is given by

$$f=4U/(E_2-E_1). \qquad \ldots\ldots(1)$$

---

* Now at Carnegie Laboratory of Physics, University College, Dundee.

If one of the fields is zero, e.g. $E_2=0$ and $E_1=E$, eqn. (1) can be written in the form $f/R=-4\delta$, where the dimensionless parameter $\delta=U/ER$ is a measure of the strength of the lens. The most representative values of δ used in practice lie between 10 and 20. The condition under which eqn. (1) was derived is $\delta \gg 1$.

Bedford (1934) discussed more fully the interpretation of eqn. (1) for finite values of δ. Gans (1937), using a numerical method, found small systematic deviations from the focal length predicted by formula (1) over the range $\delta=3{\cdot}5$ to $\delta=12$.

The present work was designed to evaluate, in a few cases, the spherical aberration constant of diverging aperture lenses about which nothing was known. It was also desired to obtain the focal length for values of δ which are no longer large.

§ 2. FIELD DISTRIBUTION WITHIN THE LENS FIELD

The focal length and spherical aberration constants were evaluated from electron trajectories traced through the lens fields. The potential distribution around an indefinitely thin diaphragm, $T=0$, perforated by a circular hole, has been calculated rigorously by Ollendorff (1932).* In the case of a diverging lens where the space to the right of the diaphragm is field-free, $E_2=0$, Ollendorff's solution for the potential along the axis can be put into the form

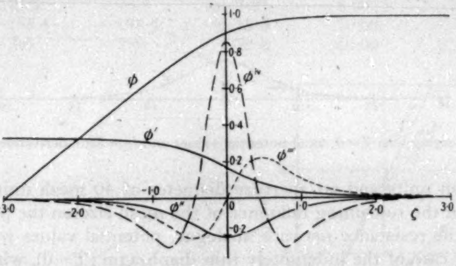


Fig. 2. Diverging lens $T=0$, axial potential values and first four derivatives ($\mu=1$, $\delta=3$).

$$\phi(\zeta)=\frac{U}{\delta}[\delta+g(\zeta)]. \qquad \ldots\ldots(2)$$

In eqn. (2) $\zeta=z/R$, the origin being at the diaphragm, and

$$g(\zeta)=\tfrac{1}{2}\zeta-(1+\zeta\tan^{-1}\zeta)/\pi. \qquad \ldots\ldots(3)$$

Equation (2) is valid under the condition $d=d_1 \gg R$, i.e. $\delta \gg 1$. In practice the deviation of Ollendorff's solution from the true potential distribution is negligible for $\delta \geqslant 3$ and still small for $\delta \simeq 2$. (The true value of $\phi(\zeta)$ is zero for $\zeta=-\delta$.)

Figure 2 shows curves of $\phi(\zeta)$ and its first four derivatives ϕ', ϕ'', ϕ''', and ϕ^{iv}, where the primes denote differentiation with respect to ζ; in fig. 2 $U=1$ and $\delta=3$.

---

* An identical solution for this problem was given by Fry (1932) and by Glaser and Henneberg (1935).

If $E_1=0$ and $E_2=E>0$, the aperture lens is convergent. Then Ollendorff's solution takes the form

$$\phi(\zeta)=U+U[\zeta-g(\zeta)]/\delta. \qquad \ldots\ldots(4)$$

Figure 3 shows values of $\phi(\zeta)$ and its first two derivatives ϕ' and ϕ'' for this case.

For lenses of finite thickness, $T\neq 0$, the potential distributions were obtained on a resistor network analogue* (Liebmann 1950). To reduce the errors due to the finite mesh size the extrapolation method described by Liebmann was used. The equivalent model has a distance between diaphragm and each anode of about 160 mesh units and an aperture diameter of 40 mesh units. To obtain an indication of the remaining influence of the mesh size on the potential values measured on the resistance network analogue, potential values were compared, in the limiting case of the indefinitely thin diaphragm ($T=0$), with Ollendorff's exact solution, eqn. (2). The extrapolated network values for $\phi(\zeta)$ were found to agree with the calculated $\phi(\zeta)$ values to within $0{\cdot}0002U$ or better. The derivatives of the axial potential distribution $\phi(z)$, needed for the computation of the trajectories, were obtained from the measured potential values by differencing in the usual manner.

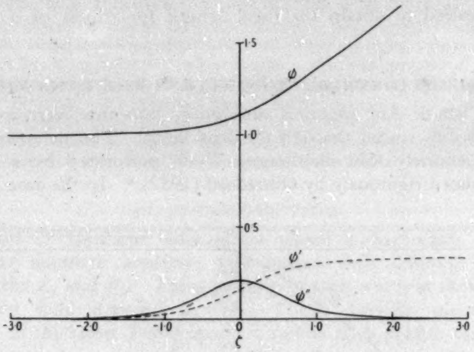


Fig. 3. Converging lens $T=0$, axial potential values and first four derivatives ($\mu=1$, $\delta=3$).

§ 3. TRAJECTORY TRACING

Paraxial trajectories were traced with the help of Liebmann's (1949) recurrence formulae

$$r_{n+1}=Q_1 r_n+Q_2 r_n', \quad r_{n+1}'=Q_3 r_n'+Q_4 r_n, \qquad \ldots\ldots(5)$$

where the Q are a power series in Δz.

For an initially parallel ray ($r_0'=0$) the focal length f in image space is given by r_0/r_1'. In a diverging lens the position z_1 of the virtual image is obtained by producing the paraxial ray back to cut the axis.

---

* This was a new precision network with 160 meshes in the z direction and 50 meshes in the r direction and increased accuracy.

The spherical aberration constant C_s can be obtained from the transverse aberration Δr_1 of the 'extra-axial' ray from the paraxial ray. The recurrence formulae are then

$$\Delta r_{n+1} = Q_1 \Delta r_n + Q_2 \Delta r_n', \quad \Delta r_{n+1}' = Q_3 \Delta r_n' + Q_4 \Delta r_n + Q_5, \quad \ldots\ldots (6)$$

where Q_5 is an expression containing the derivatives of ϕ up to the fourth order, and the r_n and r_n' values of the paraxial ray.

§4. PARAXIAL PROPERTIES OF LENSES

4.1. *Diverging Lens*, $E_1 = E$, $E_2 = 0$

(i) *Thin lens*, $T = 0$. In the limiting case of the thin diaphragm Ollendorff's potential values were used (eqn. 2). Trajectories were calculated for different values of the parameter δ, characteristic of the strength of the lens. A typical trajectory is shown in fig. 4 for $\delta = U/ER = 2$.* Table 1 gives f and z_1 for various values of δ between 2 and 100, and fig. 5 gives a plot of f as a function of δ.

Table 1. $R = 0{\cdot}1$ cm

δ	f/R	z_1/R	f(cm)	f(cm) (DC)	Diff. (%)
2	$-6{\cdot}78$	$-6{\cdot}75$	$-0{\cdot}68$	$-0{\cdot}80$	20
5	$-18{\cdot}51$	$-18{\cdot}57$	$-1{\cdot}85$	$-2{\cdot}00$	8
10	$-38{\cdot}51$	$-38{\cdot}55$	$-3{\cdot}85$	$-4{\cdot}00$	4
20	$-78{\cdot}5$	$-78{\cdot}5$	$-7{\cdot}85$	$-8{\cdot}00$	2
50	$-198{\cdot}4$	$-198{\cdot}4$	$-19{\cdot}8$	$-20{\cdot}00$	1
100	-398	-398	$-39{\cdot}8$	$-40{\cdot}00$	<1

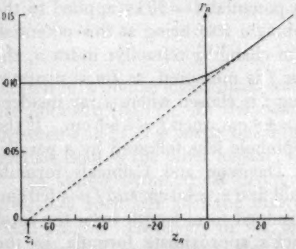


Fig. 4. Paraxial trajectory showing z_1 ($\delta = U/ER = 2$).

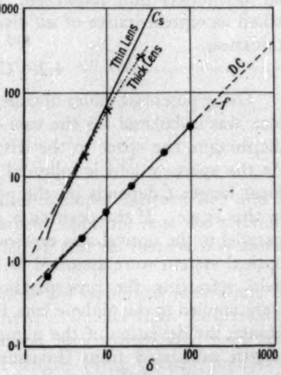


Fig. 5. Focal length and aberration constant as function of δ.

Agreement with the approximate formula, eqn. (1) (dashed line), is shown to be very close at high values of δ, i.e for 'weak lenses'; this was expected as eqn. (1) was derived under the condition $\delta \gg 1$. For lower values of δ ('strong lenses') systematic deviations from the Davisson and Calbick (DC) formula are found, the lens being always stronger than predicted by this formula; this deviation amounts to as much as 20% at $\delta=2$. The results given in fig. 5 and table 1 agree closely over the range $\delta=3\cdot5$ to 12 with those obtained by Gans (1937).

---

* Such a small value of δ was chosen to illustrate the shape of trajectories, as the refractive power of the lens is correspondingly great; this shows up better the bending of the trajectory in the lens field. Also, for values of $\delta > 2$, the virtual image point, coordinate z_1, would not be shown as $-z_1$ would be much too large.

The principal planes coincide with the centre of the lens for large values of δ. As the lens increases in strength they move slightly away from the centre by distances of the order of $0\cdot04R$.

(ii) *Lenses with finite thickness*, $T\neq0$. To study the dependence of the focal length, for given field strengths E, on the diaphragm thickness, lenses of diaphragm thickness $T=0\cdot2R$, $0\cdot4R$, $0\cdot8R$ and $1\cdot6R$ were investigated. Paraxial lens constants were determined for the relatively thick lens $T=0\cdot8R$ for the range of $\delta=5$ to 25, but only for $\delta=10$ for the other lenses. The results given in table 2 show that the focal length is almost identical with that of the lens with an indefinitely thin diaphragm, $T=0$, so that the curve shown in fig. 5 can be taken as representative of *all* divergent aperture lenses whatever the diaphragm thickness.

Table 2. f/R for $U=50$ kv, $R=0\cdot1$ cm

Thickness	20 kv/cm	50 kv/cm	100 kv/cm
0	98·4	38·6	18·5
$0\cdot2R$	—	37·9	—
$0\cdot4R$	—	38·1	—
$0\cdot8R$	98·4	38·6	18·6
$1\cdot6R$	—	38·2	—

4.2. *Converging Lens*, $T=0$.

Using potential values obtained from eqn. (4), the focal length of a converging lens was calculated for the case $\delta=5$ with a potential $U=50$ kv applied to the diaphragm, the space to the left of the diaphragm also being at this potential. As the aperture lens is followed by a space of changing refractive index n, the focal length f depends on the position where f is measured, as f/n is constant in this space. If the position z for measuring f is chosen where a ray incident parallel to the optical axis crosses the axis, $z_1=4\cdot4$ cm, then $f=+6\cdot6$ cm. If the optical system were assumed to be a 'thin' pinhole lens followed by a parallel field refracting the rays parabolically, and Davisson and Calbick's formula were applied to the pinhole lens, then one would find $z_1=4\cdot0$ cm and $f=+6\cdot0$ cm. Hence the deviation of the numerically determined focal length from the focal length calculated from Davisson and Calbick's approximate formula, in the direction of greater refracting power, is only slightly greater (10%) in this case than for the divergent lens (8%).

§ 5. SPHERICAL ABERRATION CONSTANT C_s

5.1. *Diverging Lens of Diaphragm Thickness* $T=0$

A typical plot of the deviation Δr of the 'extra-axial' ray from the paraxial ray, showing that the deviations reach a maximum near the centre of the lens, is given in fig. 6; the initial off-axis distance of this extra-axial ray incident parallel to the optical axis is $r_0/R=0\cdot1$. The spherical aberration constant C_s is defined by the relation $\Delta r_1=C_s\alpha^3$, where Δr_1 is the transverse aberration and α the angle of the ray with the optical axis. Table 3 gives the values of C_s for various values of δ from 2 to 50, and in fig. 5 C_s is plotted as a function of δ. For higher values of δ C_s follows a square law, $C_s = 1{\cdot}1\delta^2$, but for small values of δ C_s is smaller than this relation (broken line) would predict.

In order to compare the aberrations of various types of lenses and to show the connection between focal lengths and aberrations it is convenient to plot the dimensionless aberration constant $S = C_s/f$, dependent on the lens geometry, against the relative focal length f/D, where $D = 2R$. If the results obtained experimentally by Liebmann (1949) for unipotential electron lenses, and by Bruck and Romani (1944) and Ramberg (1942) using numerical methods, are plotted on a double logarithmic scale, straight lines of slopes between 1·75 and 3 are obtained (fig. 7). In the case of the divergent aperture lens the relation between δ and f/D is nearly linear (slope $=1$ in fig. 7), and the value of S is much smaller than in the other lenses.

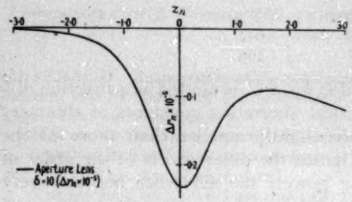


Fig. 6. Deviation of first-order ray from paraxial ray.

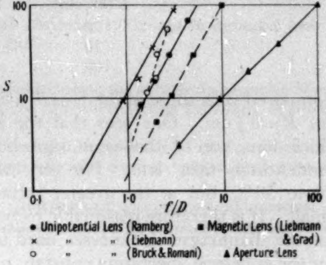


Fig. 7. Comparison of lenses.

Table 3. $R = 0{\cdot}1$ cm

δ	C_s/R	C_s (cm)	$S = C_s/f$	f/D
2	+20·7	+2·07	−3·05	−3·39
5	+177·1	+17·7	−9·57	−9·26
10	+731	+73·1	−19·0	−19·3
20	+3060	+306	−39·0	−39·2
50	+19500	+1950	−98·3	−99·2

In simple symmetrical divergent glass lenses the spherical aberration is always of such sign that the marginal focal length is smaller than the paraxial focal length. However, in the divergent electron lens investigated here the opposite has been found, the marginal focal length being greater than the paraxial focal length. In this and in the very much smaller absolute value of the spherical aberration constant C_s this divergent electron lens differs from all other electron lenses studied so far.

In view of this difference a check by Scherzer's (1936) integral formula and by Goddard's (1944) method was carried out for the case $\delta = 10$. The spherical aberration constant obtained by Scherzer's formula was identical in sign and value within a few per cent, whereas Goddard's method gave a value of the same sign but about 25% lower in value.*

---

* This appears due to the fact that in Goddard's step-by-step method the past three points are involved in every step forward. This leads to a smoothing out of the effect of the sharp peaks in the higher derivatives of the field distribution, which are mainly responsible for the aberrations, unless a greatly refined interval size is used.

Because of the 'sign' of the spherical aberration constant of the divergent aperture lens, one cannot correct the spherical aberration of a converging lens by combining it with a diverging aperture lens.

5.2. *Lens of Diaphragm Thickness $T=0{\cdot}8R$*

The spherical aberration constants for lenses of finite thickness were determined only for the very thick lens with $T=0{\cdot}8R$, as the computation of the spherical aberration constants is rather tedious. In table 4 values of C_s/R are compared with those of the lens $T=0$, and in fig. 5 C_s is plotted as a function of δ for $R=0{\cdot}1$ cm. One sees that the spherical aberration constants of this very thick lens are of the same sign but are slightly smaller than those of the 'indefinitely thin' lens. For very weak lenses the difference is of the order of 20 to 50%, but at the greater refracting powers the difference becomes very small. It is therefore reasonable to assume that over the usual working range and for the diaphragm thicknesses used in practical lenses the curve given in fig. 5 for the spherical aberration constant of the 'indefinitely thin' lens can be applied.

Table 4. C_s/R for $R=0{\cdot}1$ cm, $U=50$ kv

E (kv/cm)	$T=0$	$T=0{\cdot}8R$
20	+5000	+2750
50	+731	+642
100	+177	+166

§ 6. CONCLUSIONS

Focal lengths of aperture lenses of finite diaphragm thickness vary only very slightly with thickness, so that in practice the Davisson and Calbick 'thin lens' formula may be applied directly in the case of the diverging lens, and with a suitable modification, taking the following parallel field into account, in the case of the converging lens.

The spherical aberration constant does not seem to vary appreciably with the diaphragm thickness, particularly at relatively high refracting powers; for the diverging aperture lens met in practice it is very small and of sign opposite to that found in simple diverging glass lenses.

ACKNOWLEDGMENTS

The writer wishes to express thanks to Dr. T. E. Allibone for permission to publish this paper, to Dr. G. Liebmann for helpful suggestions and advice throughout, and to Miss E. M. Grad for help in the early stages of the work.

REFERENCES

Bedford, L. H., 1934, *Proc. Phys. Soc.*, **46,** 882.

Bruck, H., and Romani, L., 1944, *Cah. Phys.*, No. 24, 15.

Davisson, C. J., and Calbick, C. J., 1931, *Phys. Rev.*, **38,** 585; 1932, *Ibid.*, **42,** 580.

Fry, C. T., 1932, *Bell Syst. Mon.* B.671.

Gans, R., 1937, *Z. Phys.* **18,** 41.

Glaser, A., and Henneberg, W., 1935, *Z. tech. Phys.*, **16,** 229.

Goddard, L. S., 1944, *Proc. Phys. Soc.*, **56,** 372.

Liebmann, G., 1949, *Proc. Phys. Soc.* B, **62,** 213, **62,** 753; 1950, *Brit. J. Appl. Phys.*, **1,** 92.

Liebmann G., and Grad, E. M., 1951, *Proc. Phys. Soc.* B, **64,** 956.

Ollendorff, F., 1932, *Potentialfelder der Elecktrotechnik* (Berlin : Springer), p. 295.

Ramberg, E. G., 1942, *J. Appl. Phys.*, **13,** 582.

Scherzer, O., 1936, *Z. Phys.*, **101,** 593.

Investigation of Magnetic Lenses having the Axial Field $H(0, z)=\gamma/z^n$

By U. F. GIANOLA*
Department of Electron Physics, University of Birmingham

Communicated by J. Sayers ; MS. received 27th December 1951, *and in amended form 28th April* 1952

ABSTRACT. Examination of the requirements for reducing the spherical aberration of an electron lens by modifying the field shows that the hypothetical lens field $H(0, z)=\gamma/z^n$ should have small aberrations.

It is shown that the properties of a certain class of this general field configuration can be described in terms of an exact analysis. In particular, the electron optical properties of the γ/z^2 field are described in detail.

The γ/z^2 field is found to have small aberrations, but practical conditions limit appreciable improvement in resolving power. The characteristics of the γ/z^3 field are also examined and are found to give little further advantage.

.B
5
2

§ 1. INTRODUCTION

The spherical aberration of electron lenses has been the subject of extensive study, and it has been established that a positive aberration is an inherent property of lenses of the type in present use, the aberration being produced by an increase in the refractive power of the lens field with the radial distance from the optic axis (Scherzer 1936). It follows that the aberration cannot be corrected by combining existing lenses, and it remains to develop lenses which give a minimum of aberration.

In principle the spherical aberration can be reduced to any value by suitably modifying the lens field, as is shown for example by the influence that asymmetry of the axial field has upon the aberration of the magnetic lens (Dosse 1941 a). Indeed, Glaser (1940 a) has derived the form of the magnetic lens field that gives zero aberration, but the refractive power of this particular field is too small to be of practical importance (Rebsch 1940). Accordingly, it is of interest to examine the requirements for the reduction of the aberration of a practical lens and to determine the characteristics of a suitable lens field.

The approach to this problem is indicated by an examination of the general expression giving the spherical aberration in third-order approximation (Glaser 1940 a), namely

$$\Delta_s = \frac{r_X'}{\phi_X^{1/2}} \int_X^{z_1} [L r_\alpha^4 + 2M r_\alpha^2 r_\alpha'^2 + N r_\alpha'^4]\, dz,$$

where r_X' is the slope of the trajectory at the object point $z=X$, and $r_\alpha(z)$ is the specific paraxial trajectory that originates in the object point with unit inclination to the optic axis. L, M, and N are functions of the axial distributions of the electric and magnetic components of the lens field, $\phi(0,z)$ and $H(0,z)$ respectively, and their derivatives, i.e.

---

* Now at University of British Columbia, Canada.

$$L=\frac{\phi^{-1/2}}{32}\left[\phi^{-1}\left(\phi''+\frac{e}{m}H^2\right)^2-2\frac{e}{m}HH''-\phi^{\mathrm{iv}}\right],$$

$$N=\tfrac{1}{2}\phi^{1/2},\qquad M=\frac{\phi^{-1/2}}{8}\left[\phi''+\frac{e}{2m}H^2\right].$$

Many of the variables entering into this expression are interdependent and it is difficult to determine a practical approach for minimizing the aberration, except to specify that a strong negative term in the integral is required. For a purely electrostatic lens this requirement is partially fulfilled if the fourth and second derivatives of the field are respectively positive and negative in the region between the object and the image. In a continuous lens field these two requirements are satisfied mutually only over a limited region about the maxima of the distribution. Thus, this condition, together with the simultaneous requirement for a rapid convergence of the electron trajectory, indicates that the optimum field is one that includes a sharp maximum, with the object situated in that region of the field which is varying rapidly. Unfortunately, the field gradient is limited in practice by auto-electronic emission from the electrodes, and it is also undesirable to locate the object in a region of high field strength. It appears that there is little improvement to be obtained in this manner.

The corresponding condition for the aberration of a purely magnetic lens to be a minimum requires that the second derivative of the field be highly positive in the region between object and image. This is also a necessary condition for the aberration of a composite electric and magnetic lens to be a minimum, the consideration given to a purely electrostatic lens also applying. These conclusions are in general agreement with those given by Rebsch (1938), who showed that the aberration may be reduced by immersing the object in a more rapidly varying field. It should be stressed that application of these relations is very limited by practical restrictions to the field intensities. Nevertheless, some small improvement is possible. These general conclusions will be considered in relation to the purely magnetic lens in order to ascertain the limit to practical improvement.

The field of a typical magnetic lens is approximated to by the analytical field $H(0,z)=H(0,0)/\{1+(z/a)^2\}$, and it is often convenient to discuss its performance in terms of this analytical field (Glaser 1941, Dosse 1941 b, Hesse 1950). The second derivative of this axial field is positive only for values $(z/a)^2>\frac{1}{3}$. Thus, at the centre of the lens where maximum refraction occurs, the second derivative is negative so that a positive contribution to the basic spherical aberration results. This is indicated by the corresponding influence of the object position (Glaser 1941).

Consider now the hypothetical field having the general form $H(0,z)\propto 1/z^n$. The second derivative of this axial field is always positive. Accordingly, this lens field should give a smaller aberration than does the $1/\{1+(z/a)^2\}$ field; for this reason its characteristics have been examined in detail. In practice this field cannot be reproduced completely because of the discontinuity. However, a part of the field can be produced by a suitable choice of lens geometry. Thus, if the object is regarded as being self-luminous, as is often permissible, the characteristics of the lens may be determined by consideration of the analytical field alone if the object be placed in that part of the field approximating to the theoretical distribution, provided that the entrance aperture can be suitably adjusted. A description of the properties of the $1/z^n$ magnetic lens field forms the content of the remainder of this paper.

§ 2. ELECTRON MOTION IN THE $1/z^n$ MAGNETIC LENS FIELD

An analytical solution to the electron trajectory in a lens field has been obtained for only a few lens fields (Glaser 1941, 1949, Hutter 1945, Rüdenberg 1949). Therefore, it is interesting to find that a certain class of the general $1/z^n$ lens field is amenable to exact analysis in terms of elementary quantities.

The equation of motion of a paraxial electron in the magnetic field $H(0, z)=\gamma/z^n$, where γ is a constant characteristic of the lens, may be expressed

$$r''+B^2r/z^{2n}=0, \qquad \ldots\ldots(1)$$

where $B=(e\gamma^2/8m\phi_X)^{1/2}$ is a constant which determines the power of the lens, and has the dimension of length alone. By making the substitution $u=r'/r$, the ray equation may be reduced to a differential equation of the first order in the form of the Riccati equation, and is integrable in finite terms if

$$n=2g/(2g\pm 1), \qquad \ldots\ldots(2)$$

where g is zero or a positive integer (Forsyth 1929).

Thus, an analytical solution to the paraxial electron trajectory can be obtained for fields that are defined by eqn. (2). Of these the $1/z^2$ field is of particular interest since it has the strongest divergence. This lens field may be approximated to by a lens having the geometry illustrated by fig. 1, with the object mounted in the divergent part of the field, this arrangement allowing the object to be situated in a region of high field intensity. The electron imaging properties of this particular lens field are examined in the following section.

§ 3. THE DIOPTRICS OF THE $1/z^2$ LENS FIELD

The trajectory of a paraxial electron in the $1/z^2$ lens has the solution

$$r=Dz((1+K)\cos(B/z)+i(1-K)\sin(B/z)); \qquad \ldots\ldots(3)$$

$D(1+K)$ and $D(1-K)$ are constants of integration determined by the initial conditions that specify a particular trajectory. The location of the gaussian image plane is given by the point of intersection with the optic axis of the specific paraxial ray $r_\alpha(z)$. Equation (3) gives for this trajectory $r_\alpha(z)=-(z/\omega)\sin\psi$, where $\omega=B/X$, $\psi=B(z^{-1}-X^{-1})$. It follows that paraxial trajectories intersect the optic axis at points given by the relations

$$z=0, \quad \psi=0, \ \pm\pi, \ \pm 2\pi, \ldots, \ \pm N\pi. \qquad \ldots\ldots(4)$$

This value of z and the positive values of ψ represent the infinite number of virtual intersections of the trajectory with the axis in the region preceding the object point. Hence these values do not influence practical considerations and will be neglected. The N negative values of ψ correspond to N successive image points. Consequently the condition for single image formation may be expressed $2\pi>\omega\geqslant\pi$. Accordingly, the position of the gaussian image plane under conditions of single image formation is given by the expression $z_1=B/(\omega-\pi)$. The corresponding magnification of the gaussian image is given by the relation $m=\omega/(\omega-\pi)$ where ω is again to satisfy the condition expressed for single image formation.

It will be observed that when $\omega=\pi$ the image is formed at infinite distance with infinite magnification, the corresponding object point $X_f=B/\pi$ being the effective focal point of the lens. Accordingly, the effective focal length of the lens is given by the relation $f=\lim_{z\to\infty} r_f(z)$ where $r_f(z)$ is the paraxial ray that originates in the focal point with unit inclination to the axis. Accordingly, eqn. (3) gives for the effective focal length $f=B/\pi$, giving a simple measure of the refractive power of the lens.

The general gaussian properties of the lens are now determined explicitly in terms of the field parameters. It remains to evaluate the aberrations.

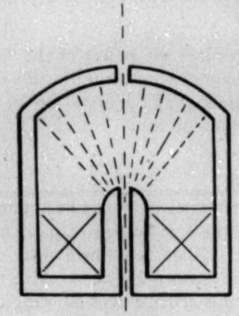


Fig. 1. The diagrammatic cross section of a magnetic lens which is suitable for reproducing a part of the hypothetical field distribution $H(0, z)=\gamma/z^2$.

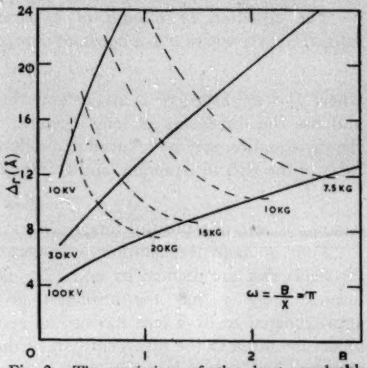


Fig. 2. The variation of the least resolvable distance Δr with the lens power parameter B of the $1/z^2$ magnetic lens field at specific electron energies. The power of the lens is maintained at the high magnification condition throughout, and the loci of the field strengths that are appropriate to this condition are given in kilogauss.

§ 4. THE RESOLVING POWER OF THE $1/z^2$ MAGNETIC LENS FIELD

For this particular lens field the spherical aberration, defined above in third order approximation, becomes

$$\frac{\Delta_s}{r_X'^3}=C_s=\frac{B}{2\omega^4}\left[\omega(\omega-\pi)+\frac{\pi^2}{3}+\frac{1}{4}\right]. \qquad \ldots\ldots(5)$$

In addition to the spherical aberration, a chromatic aberration produced by fluctuations $e\Delta\phi$ in the energy of the electron beam is important in considerations of resolving power. For a purely magnetic lens this aberration may be expressed, in third order approximation (Glaser 1940 b),

$$\Delta_c=\frac{\Delta\phi}{\phi}\, r_X' \int_X^{z_1} r_\alpha'^2\, dz, \qquad \ldots\ldots(6)$$

where $r_\alpha(z)$ is again the trajectory that originates in the object point with unit inclination to the axis, with the added condition that it vanishes in both object and image points. For the $1/z^2$ lens field this aberration may be expressed

$$\Delta_c/r_X'=C_c\Delta\phi/\phi=(\pi/2B)X^2\Delta\phi/\phi.$$

A comparison of the quality of different electron lenses is difficult because of the many independent parameters involved, and it is customary to take the dimensionless constants C_s/f and C_c/f as characteristic figures representing the aberrations produced by lenses of equal focal length. For the $1/z^2$ lens field with the object maintained at the high magnification position, these characteristic figures take the values $(C_s/f)_{X_f} = 0{\cdot}18$, $(C_c/f)_{X_f} = 0{\cdot}5$. The corresponding high magnification characteristics of the $1/\{1 + (z/a)^2\}$ lens field are dependent upon the parameter k^2 (Glaser 1941), but for a lens power maintained at the optimum value for a specific voltage (Cosslett 1946) the aberration characteristics take the values $(C/f_s)_{X_f} = 0{\cdot}96$, $(C_c/f)_{X_f} = 0{\cdot}76$. In terms of these characteristics alone the $1/z^2$ lens field gives a marked improvement. However, little indication is given of any improvement in resolving power that may be obtained in practice, bearing in mind the limitation set by magnetic saturation of the pole pieces. A more satisfactory comparison is given by the limit of resolution obtained under high magnification conditions with the lens aperture adjusted to equalize the discs of confusion produced by spherical aberration and by diffraction (Zworykin *et al.* 1945). This limit Δ_r is plotted for the $1/z^2$ lens in fig. 2, in dependence upon the power parameter B for specific values of the accelerating voltage ϕ. The contribution of the chromatic aberration is calculated for a voltage ripple of $2{\cdot}5 \times 10^{-5}$, and a spread of one electron volt in the energy of the imaging beam. The loci of the appropriate field strengths at the object, $H(0, X_f)$, that are necessary to maintain the high magnification condition are also shown.

A comparison of the resolving power of this lens field with that given by the $1/\{1 + (z/a)^2\}$ field (Cosslett 1946, 1947, Gianola 1950) shows that it is little different. It is not possible to ascertain the absolute ultimate resolving powers of the two lens fields without some knowledge of the practical field intensities that are feasible, but it is expected that the peak field strength of the $1/z^2$ lens field may approach saturation field strength, since the object may be placed close to the pole pieces. Consequently some small improvement in resolving power may be possible.

§ 5. FURTHER INVESTIGATION OF THE GENERAL $1/z^n$ LENS FIELD

It follows from eqn. (2) that the fields that can be analysed in the manner described all have a field divergence equal to or less than that of the $1/z^2$ lens. The spherical aberration of the latter field was not expected to differ appreciably from that of the $1/\{1 + (z/a)^2\}$ lens field with an infinite asymmetry index, $q = \infty$ (Dosse 1941 a). For this reason the performance of a rather more divergent field has been examined.

Referring again to the paraxial ray equation, eqn. (1), and making the substitution $u = 1/z$, $p = 1/2(n-1)$, the ray equation is reduced to the form of the general Bessel equation. Accordingly, for non-integer values of p the trajectory may be expressed in terms of Bessel functions of the first kind, namely

$$r = u^{-1/2}(A_1 J_p(x) + A_2 J_{-p}(x)), \qquad \ldots\ldots(7)$$

where A_1 and A_2 are integration constants and $x = 2Bpu^{1/2p}$. The position of the gaussian image plane can be determined as previously by finding the axial intersections of the specific trajectory $r_\alpha(z)$; this gives the general relation $J_p(x_1)J_{-p}(x_X) = J_p(x_X)J_{-p}(x_1)$ where x_X and x_1 represent the values of x at the object and image points respectively.

Further, the requirement that the trajectory should approach zero slope at infinite distance gives for the focal point the relation $J_p(x_f)=0$. The effective focal length of the general $1/z^n$ field may now be obtained as previously, and can be expressed in the form

$$\frac{f}{X_f} = -\left(\frac{x_f}{2}\right)^p \frac{p\pi}{\pi(p)\sin p\pi} J_{-p}(x_f),$$

where $\pi(n)$ is the gauss pi-function.

It remains finally to obtain an explicit relation for the spherical aberration that is produced by the general $1/z^n$ lens field. Substituting for the axial field $H(0, z)=\gamma/z^n$ and for the paraxial trajectory, eqn. (7), into the expression for the aberration constant

$$\frac{C_s}{X} = \frac{-(x_X)^{-1/(n-1)}}{12(n-1)} \left(\frac{\pi}{\sin p\pi}\right)^4 \int_{x_X}^{x_i} \left[x^2 + \frac{n(4n-1)}{16(n-1)^2}\right] \times [J_p(x_X)J_{-p}(x) - J_{-p}(x_X)J_p(x)]^4\, x^{n/(n-1)}\, dx. \qquad \ldots\ldots(8)$$

It is now possible to determine the main characteristics of a specific example of the general $1/z^n$ lens field, given tabulated values of the appropriate Bessel function. For the particular instance of the $1/z^2$ lens field, $p=\frac{1}{2}$, this general analysis yields expressions which correspond to those obtained by direct integration of the ray equation.

In particular the $1/z^3$ lens field has been investigated, it being anticipated that distributions of greater divergence would have too small a refractive power to be of importance. For this distribution, $p=\frac{1}{4}$, an inspection of tables of $J_{\pm 1/4}(x)$ (National Bureau of Standards 1949), gives for the condition of single image formation $5{\cdot}91 > B/2X^2 \geqslant 2{\cdot}79$ and for the position of the first effective focal point $x_f = 2{\cdot}79$. In addition, a numerical integration of eqn. (8) between the limits $x_X = 2{\cdot}79$ and $x_i = 0$, gives for the aberration constant of the $1/z^3$ lens field, under conditions of high magnification, the value $C_s/X_f = 0{\cdot}10$.

§ 6. COMPARISON OF EXAMPLES OF THE $1/z^n$ LENS FIELDS

The quantity $C_s H(0, X_f)$ is best taken as a characteristic figure in the comparison of $1/z^n$ lens fields, as it gives a measure of the aberrations produced by fields having objects situated at points having comparable field strengths. The table gives this aberration characteristic, together with the effective focal point and focal length characteristics, for the $1/z^{4/3}$, $1/z^2$, and $1/z^3$ lenses, to show the variation over a practical range.

n	4/3	2	3
X_f	4·49	3·14	2·79
$(e/8m\phi_X)^{1/2}\, X_f H(0, X_f)$	1·50	3·14	5·58
$(e/8m\phi_X)^{1/2}\, f H(0, X_f)$	27·6	3·14	2·52
$(e/8m\phi_X)^{1/2}\, C_s H(0, X_f)$	2·6	0·57	0·56

It is apparent that little practical improvement in the quality of the lens is obtained by increasing the divergence of the field beyond that of the $1/z^2$ lens. For this reason the resolving power of fields of high divergence will not be considered further.

This analysis gives some indication of the lower limit that the spherical aberration may be reduced to in practice by a straightforward modification of the field of an electron lens. It is concluded from the results of the investigation that, in the case of the magnetic lens at least, no outstanding improvement is to be obtained in this way without introducing subsidiary corrector fields.

The fact that the $1/z^n$ lens fields can be subjected to rigorous analysis facilitates their use in design problems, in particular those concerned with the investigation of suitable corrector lenses, when their low aberration characteristics may prove advantageous.

ACKNOWLEDGMENTS

I am indebted to the Department of Scientific and Industrial Research for its support while engaged in this work, and to Professor J. Sayers for his interest and advice.

REFERENCES

Cosslett, V. E., 1946, *Proc. Phys. Soc.*, **58**, 443; 1947, *J. Sci. Instrum.*, **24**, 40.
Dosse, J., 1941 a, *Z. Phys.*, **117**, 316; 1941 b, *Ibid.*, **117**, 437.
Forsyth, A. R., 1929, *A Treatise on Differential Equations* (London: Macmillan, 6th edn.), p. 191.
Gianola, U. F., 1950, *Proc. Phys. Soc.* B, **63**, 703.
Glaser, W., 1940 a, *Z. Phys.*, **116**, 19; 1940 b, *Ibid.*, **116**, 56; 1941, *Ibid.*, **117**, 285; 1949, *Acta Phys. Austriaca*, **3**, 38.
Hesse, M. B., 1950, *Proc. Phys. Soc.* B, **63**, 386.
Hutter, R. G., 1945, *J. Appl. Phys.*, **16**, 678.
National Bureau of Standards, 1949, *Tables of Bessel Functions of Fractional Order* (Columbia: University Press).
Rebsch, R., 1938, *Ann. Phys., Lpz.*, **31**, 551; 1940, *Z. Phys.*, **116**, 729.
Rüdenberg, R., 1949, *J. Franklin Inst.*, **246**, 311, 377.
Scherzer, O., 1936, *Z. Phys.*, **101**, 593.
Zworykin, V. K., *et al.*, 1945, *Electron Optics and the Electron Microscope* (New York: Wiley), p. 699.

Analysis of Polarized Light Reflected from Absorbing Materials at Normal Incidence

By J. WOODROW, B. W. MOTT and H. R. HAINES
Atomic Energy Research Establishment, Harwell, Didcot, Berks.

Communicated by L. F. Bates; MS. received 6th *March* 1952

ABSTRACT. Expressions are derived for the intensity of the analysed light reflected from absorbing crystals (*a*) under crossed polarizing units and (*b*) with the analyser slightly rotated from the crossed position. Four intensity maxima per complete rotation of the specimen are obtained when the polarizers are crossed, but under certain circumstances, as the analyser is rotated from the crossed position, the number of maxima may be reduced to two. These relationships are illustrated by experimental curves on specimens of tin, uranium, bismuth and stibnite.

§ 1. INTRODUCTION

In an earlier paper (Mott and Haines 1951 a) the effect of crystal symmetry on the reflection of a plane polarized beam of light at normal incidence was summarized, together with the modifications produced by anisotropic surface films and controlled etching of the reflecting surface.

According to the theory developed by various authors (Szivessy 1929, Wright 1919, Orcel 1925, Capdecomme and Orcel 1941), four equal intensity maxima and minima should be obtained when an anisotropic specimen is rotated over a range of azimuth angles from 0°–360° between crossed polarizers. This is confirmed by observation in some cases, but many anomalies have been reported. Farnham (1931) and Short (1940) classify anisotropic minerals according to whether they give four or two intensity maxima when the crystal is rotated through 360°. Eash and Upthegrove (1933) state that the anisotropic δ' constituent in ternary copper–nickel–tin alloys gives two extinctions per complete stage rotation. Sampson (1923) measured the rotation angle produced by reflection from various faces of a manganite crystal, unequal maximum rotations in the positive and negative directions being obtained. In addition, he found that the angle between successive intensity maxima and also between successive minima was not exactly 90° as would be expected. Sampson (1929) subsequently amplified his results and showed that as the analyser is rotated from the crossed position intermediate maxima are reduced in intensity until finally only two maxima are obtained when the offset of the analyser is equal to the maximum angle of rotation of the plane of polarization. The maximum contrast between the grains however is obtained when the analyser is rotated still further. Orcel (1925) also points out that, when the effect of reflection is essentially a rotation of the plane of polarization (i.e. the ellipticity of the reflected beam is very slight), the contrast between the grains in a multi-grained specimen is most marked when either the polarizer or analyser is rotated slightly from the crossed position. Jones (1924) found that most etched metals gave four maxima, but with bismuth lightly etched in nitric acid a marked difference in the alternate pairs of maxima can be detected when the analyser is rotated only slightly from the crossed position. Jones suggests that this affords an accurate method of crossing the polarizer and analyser. On the other hand, Morrogh (1949), for example, reports that the anisotropic η-phase of the iron–silicon system gives four maxima.

Variable results were obtained in the earlier work at A.E.R.E., especially on electrolytically polished uranium, although some specimens always gave four intensity maxima per rotation. For instance, in a given field of a uranium specimen, some grains gave two intensity maxima, others gave four, and only in a few cases were the four maxima of equal intensity, two being more marked than the other pair. The maximum grain contrast for uranium and many other anisotropic metals as polished was obtained when the analyser was rotated through 1°–2° from the crossed position. When a Foster prism (1938) was used as the vertical illuminator, the results were generally similar to those with polaroid units and a glass slip vertical reflector, although when four maxima were obtained they appeared to be more equal in intensity. If a quarter-wave plate was inserted between the Foster prism and the objective, it was often possible to rotate the plate to such a position that four nearly equal maxima were obtained for most grains on a particular specimen. The setting of the quarter-wave plate varied with the specimen under examination and the objective lens used.

In view of these anomalies it was decided to investigate the phenomenon both in theory and in practice. The intensity measurements were made with the photometer unit previously described placed in the monocular eyepiece tube of a Beck Universal Metallograph No. 50 (Mott and Haines 1951 a).

§2. EXPRESSIONS OF PRACTICAL INTEREST

2.1. *Low Symmetry Crystals*

2.11. *General case for low symmetry crystals* (i.e. triclinic, monoclinic and orthorhombic systems). Whilst transparent crystals of these three systems are optically biaxial, absorbing crystals with the same symmetry do not possess true optic axes. For a given direction of the wave normal, such crystals in general transmit two elliptically polarized waves. The ellipses representing the polarization of the two waves have the same ellipticity, and are described in the same sense; the major axes are mutually perpendicular.

A normally incident wave reaching the surface of an absorbing crystal enters it without any change in the direction of the wave normal. Two principal elliptically polarized waves can therefore be associated with each case of normal incidence, and will vary with the orientation of the surface and crystal symmetry. These two principal refracted waves are transmitted with different velocities, and are absorbed to different extents; associated with them are two principal reflected waves with different reflectivities and phase shifts.

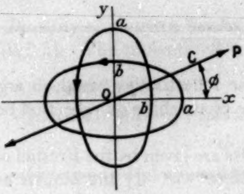


Fig. 1.

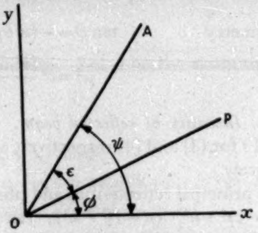


Fig. 2.

Only a principal incident beam can be transmitted and reflected without any change in polarization. In dealing with a plane polarized incident beam it is first necessary to resolve it into two principal elliptically polarized beams.

2.12. *Resolution of plane-polarized normally-incident beam into two principal elliptically polarized beams.* The axes of coordinates for the surface are most conveniently taken to coincide with the axes of the principal ellipses of polarization for a wave normal perpendicular to the surface.

Let the azimuth of the plane of polarization of the normally incident beam be ϕ and let the amplitude of vibration be c (fig. 1). The vibration representing the plane polarized beam is

$$x = c\cos\phi\cos\theta; \quad y = c\sin\phi\cos\theta \qquad \ldots\ldots(1)$$

where $\theta = 2\pi ft$, f being the light frequency.

The two principal elliptically polarized waves that can be transmitted without change of polarization correspond to vibrations with loci

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\ ; \quad \frac{x^2}{b^2}+\frac{y^2}{a^2}=1. \qquad \ldots\ldots(2)$$

The vibrations corresponding to the two elliptically polarized beams are therefore

$$x_1 = a\cos(\theta+\alpha)\ ; \quad y_1 = b\sin(\theta+\alpha) \qquad \ldots\ldots(3)$$

and

$$x_2 = b\cos(\theta+\beta)\ ; \quad y_2 = a\sin(\theta+\beta), \qquad \ldots\ldots(4)$$

where α and β are phase angles relative to the incident plane polarized beam given by eqn. (1).

Let vibration $(1) \equiv \lambda \times$ vibration $(3) + \mu \times$ vibration (4) so that

$$c\cos\phi\cos\theta \equiv \lambda a\cos(\theta+\alpha)+\mu b\cos(\theta+\beta),$$
$$c\sin\phi\cos\theta \equiv \lambda b\sin(\theta+\alpha)+\mu a\sin(\theta+\beta).$$

By equating coefficients of $\cos\theta$ and $\sin\theta$ we find for the unknowns α, β, λ, μ

$$\lambda a\cos\alpha+\mu b\cos\beta=c\cos\phi, \qquad \lambda a\sin\alpha+\mu b\sin\beta=0,$$
$$\lambda b\cos\alpha+\mu a\cos\beta=0, \qquad \lambda b\sin\alpha+\mu a\sin\beta=c\sin\phi.$$

Eliminating μ and β,

$$\lambda(a^2-b^2)\cos\alpha=ac\cos\phi, \quad \lambda(a^2-b^2)\sin\alpha=-bc\sin\phi, \qquad \text{......(5)}$$

whence $\tan\alpha=-(b/a)\tan\phi,$

$$\lambda^2=c^2\frac{(a^2\cos^2\phi+b^2\sin^2\phi)}{(a^2-b^2)^2}. \qquad \text{......(6)}$$

By symmetry $\tan\beta=-(a/b)\tan\phi$

$$\mu^2=c^2\frac{(b^2\cos^2\phi+a^2\sin^2\phi)}{(a^2-b^2)^2}. \qquad \text{......(7)}$$

2.13. *Intensity of reflected beam.* Let the reflectivities based on amplitudes be r and s for (3) and (4) respectively; let the phase shifts on reflection be γ and δ respectively.

The principal reflectivities and phase shifts are given by the Fresnel equations $r^2=\{(n-1)^2+k^2\}/\{(n+1)^2+k^2\}$; $\tan\gamma=-2k/(k^2+n^2-1)$ and similar equations for s and δ; n and k denote the principal refractive indices and principal absorption coefficients for the two transmitted principal elliptically polarized waves.

The reflected wave is given by

$$x_R=r\lambda a\cos(\theta+\alpha+\gamma)+s\mu b\cos(\theta+\beta+\delta),$$
$$y_R=r\lambda b\sin(\theta+\alpha+\gamma)+s\mu a\sin(\theta+\beta+\delta).$$

The intensity of the reflected (and unanalysed beam) is therefore

$$\begin{aligned} I_R=&\{r\lambda a\cos(\alpha+\gamma)+s\mu b\cos(\beta+\delta)\}^2+\{r\lambda a\sin(\alpha+\gamma)+s\mu b\sin(\beta+\delta)\}^2\\ &+\{r\lambda b\sin(\alpha+\gamma)+s\mu a\sin(\beta+\delta)\}^2+\{r\lambda b\cos(\alpha+\gamma)+s\mu a\cos(\beta+\delta)\}^2\\ =&(a^2+b^2)(r^2\lambda^2+s^2\mu^2)+4rs\lambda\mu ab\cos(\alpha-\beta+\gamma-\delta). \qquad \text{......(8)}\end{aligned}$$

λ^2 and μ^2 are given by eqns. (6) and (7), $\lambda\cos\alpha$ and $\lambda\sin\alpha$ are given by eqn. (5); by symmetry we also have

$$\mu(a^2-b^2)\cos\beta=-bc\cos\phi, \quad \mu(a^2-b^2)\sin\beta=ac\sin\phi. \qquad \text{......(5 a)}$$

With these substitutions eqn. (8) becomes

$$\begin{aligned}\frac{(a^2-b^2)^2}{c^2}I_R=&(a^2+b^2)(r^2a^2\cos^2\phi+r^2b^2\sin^2\phi+s^2b^2\cos^2\phi+s^2a^2\sin^2\phi)\\ &+4rsab\{-ab\cos(\gamma-\delta)+(a^2-b^2)\sin\phi\cos\phi\sin(\gamma-\delta)\}\\ =&\tfrac{1}{2}(r^2+s^2)(a^2+b^2)^2-4rsa^2b^2\cos(\gamma-\delta)+\tfrac{1}{2}(r^2-s^2)(a^4-b^4)\cos 2\phi\\ &+2rsab(a^2-b^2)\sin(\gamma-\delta)\sin 2\phi.\end{aligned}$$

Let the ellipticity b/a of the principal vibrations corresponding to the section be e. Then

$$\begin{aligned}\frac{I_{\mathrm{R}}}{c^2}=\frac{(r^2+s^2)}{2}\,\frac{(1+e^2)^2}{(1-e^2)^2}-4rs\frac{e^2}{(1-e^2)^2}\cos(\gamma-\delta)+\frac{(r^2-s^2)}{2}\,\frac{1+e^2}{1-e^2}\cos 2\phi&\\+2rs\frac{e}{1-e^2}\sin(\gamma-\delta)\sin 2\phi.&\qquad\ldots\ldots(9)\end{aligned}$$

As the plane of polarization is rotated through 360°, the intensity I_{R} will pass through two maxima and two minima. These points are given by

$$\tan 2\phi_0=\frac{4rs}{r^2-s^2}\,\frac{e}{1+e^2}\sin(\gamma-\delta)=\frac{4Ke}{(K^2-1)(1+e^2)}\sin(\gamma-\delta),\qquad\ldots\ldots(10)$$

where K is the ratio of principal reflectivities r/s.

Hence the directions of maximum and minimum reflectivity do *not* in general lie along the principal directions of the section. These directions coincide if the ellipticity e is zero as for uniaxial crystals, or if the phase difference $\gamma-\delta$ happens to be zero.

2.14. *Intensity of analysed beam (general case).* Let ψ be the azimuth of the analyser (fig. 2). The transmitted wave is

$$\begin{aligned}x_{\mathrm{A}}&=x_{\mathrm{R}}\cos\psi+y_{\mathrm{R}}\sin\psi\\&=\{r\lambda a\cos(\theta+\alpha+\gamma)+s\mu b\cos(\theta+\beta+\delta)\}\cos\psi\\&\qquad+\{r\lambda b\sin(\theta+\alpha+\gamma)+s\mu a\sin(\theta+\beta+\delta)\}\sin\psi\\&=[\{r\lambda a\cos(\alpha+\gamma)+s\mu b\cos(\beta+\delta)\}\cos\psi+\{r\lambda b\sin(\alpha+\gamma)\\&\qquad+s\mu a\sin(\beta+\delta)\}\sin\psi]\cos\theta+[-\{r\lambda a\sin(\alpha+\gamma)\\&\qquad+s\mu b\sin(\beta+\delta)\}\cos\psi+\{r\lambda b\cos(\alpha+\gamma)\\&\qquad+s\mu a\cos(\beta+\delta)\}\sin\psi]\sin\theta.\end{aligned}$$

The intensity of the transmitted beam is therefore

$$\begin{aligned}I&=[\{r\lambda a\cos(\alpha+\gamma)+s\mu b\cos(\beta+\delta)\}^2+\{r\lambda a\sin(\alpha+\gamma)+s\mu b\sin(\beta+\delta)\}^2]\cos^2\psi\\&\qquad+[\{r\lambda b\sin(\alpha+\gamma)+s\mu a\sin(\beta+\delta)\}^2+\{r\lambda b\cos(\alpha+\gamma)\\&\qquad+s\mu a\cos(\beta+\delta)\}^2]\sin^2\psi+2[\{r\lambda a\cos(\alpha+\gamma)\\&\qquad+s\mu b\cos(\beta+\delta)\}\{r\lambda b\sin(\alpha+\gamma)+s\mu a\sin(\beta+\delta)\}\\&\qquad-\{r\lambda a\sin(\alpha+\gamma)+s\mu b\sin(\beta+\delta)\}\{r\lambda b\cos(\alpha+\gamma)\\&\qquad+s\mu a\cos(\beta+\delta)\}]\sin\psi\cos\psi\\&=(r^2\lambda^2a^2+s^2\mu^2b^2)\cos^2\psi+(r^2\lambda^2b^2+s^2\mu^2a^2)\sin^2\psi\\&\qquad-2rs\lambda\mu(a^2-b^2)\sin(\alpha-\beta+\gamma-\delta)\sin\psi\cos\psi+2rs\lambda\mu ab\cos(\alpha-\beta+\gamma-\delta).\end{aligned}$$

Substituting from eqns. (6), (7), (5) and (5 *a*), we obtain

$$\begin{aligned}(a^2-b^2)^2I/c^2&=\{(r^2a^4+s^2b^4)\cos^2\phi+(r^2+s^2)a^2b^2\sin^2\phi\}\cos^2\psi+\{(r^2+s^2)a^2b^2\cos^2\phi\\&\qquad+(r^2b^4+s^2a^4)\sin^2\phi\}\sin^2\psi+2rs(a^2-b^2)\{(a^2-b^2)\cos\phi\sin\phi\cos(\gamma-\delta)\\&\qquad+ab\sin(\gamma-\delta)\}\sin\psi\cos\psi+2rsab\{-ab\cos(\gamma-\delta)\\&\qquad+(a^2-b^2)\cos\phi\sin\phi\sin(\gamma-\delta)\}.\end{aligned}$$

Polarizer and crystal fixed; *analyser rotated.* In this case ϕ is constant and ψ varies; expressing the intensity as a sum of harmonics in ψ, we obtain

$$\begin{aligned}(a^2-b^2)^2 I/c^2 = {} & \tfrac{1}{2}(a^2+b^2)\{(r^2a^2+s^2b^2)\cos^2\phi+(r^2b^2+s^2a^2)\sin^2\phi\} \\ & +2rsab\{(a^2-b^2)\cos\phi\sin\phi\sin(\gamma-\delta)-ab\cos(\gamma-\delta)\} \\ & +\tfrac{1}{2}(a^2-b^2)[\{(r^2a^2-s^2b^2)\cos^2\phi+(r^2b^2-s^2a^2)\sin^2\phi\}^2 \\ & +4r^2s^2\{(a^2-b^2)\cos\phi\sin\phi\cos(\gamma-\delta)+ab\sin(\gamma-\delta)\}^2]^{1/2}\cos 2(\psi-\psi_0),\end{aligned} \qquad \ldots\ldots(11)$$

where the phase angle ψ_0 is given by

$$\tan 2\psi_0 = \frac{2rs\{(a^2-b^2)\cos\phi\sin\phi\cos(\gamma-\delta)+ab\sin(\gamma-\delta)\}}{(r^2a^2-s^2b^2)\cos^2\phi+(r^2b^2-s^2a^2)\sin^2\phi},$$

i.e.
$$\tan 2\psi_0 = 2K\,\frac{(1-e^2)\cos\phi\sin\phi\cos(\gamma-\delta)+e\sin(\gamma-\delta)}{(K^2-e^2)\cos^2\phi+(K^2e^2-1)\sin^2\phi}. \qquad \ldots\ldots(11\,a)$$

Maximum illumination occurs for analyser setting ψ_0 and $\psi_0+\pi$; minimum illumination occurs at $\psi_0+\pi/2$ and $\psi_0+3\pi/2$. Thus two maxima occur per revolution of the analyser.

When $\phi=0$, i.e. when polarizer is set parallel to principal direction $\tan 2\psi_0 = 2Ke/(K^2-e^2)\sin(\gamma-\delta)$. Hence the maximum illumination does not, in the general case, occur when the polarizer and analyser are parallel to a principal direction of the section.

Polarizer and analyser fixed; *crystal rotated.* If the polarizer and analyser are set at angle ϵ, $\psi=\phi+\epsilon$. In this case both ϕ and ψ vary but ϵ remains constant. From eqn. (11)

$$\begin{aligned}(a^2-b^2)^2 I/c^2 = {} & \tfrac{1}{4}(a^2+b^2)^2(r^2+s^2)-2rsa^2b^2\cos(\gamma-\delta) \\ & +\tfrac{1}{4}(a^4-b^4)(r^2-s^2)(\cos 2\phi+\cos 2\psi)+rsab(a^2-b^2)\sin(\gamma-\delta)(\sin 2\phi+\sin 2\psi) \\ & +\tfrac{1}{4}(a^2-b^2)^2(r^2+s^2)\cos 2\phi\cos 2\psi+\tfrac{1}{2}rs(a^2-b^2)^2\cos(\gamma-\delta)\sin 2\phi\sin 2\psi \\ = {} & \tfrac{1}{4}(a^2+b^2)^2(r^2+s^2)-2rsa^2b^2\cos(\gamma-\delta)+\tfrac{1}{8}(a^2-b^2)^2\{r^2+s^2+2rs\cos(\gamma-\delta)\}\cos 2\epsilon \\ & +\tfrac{1}{2}(a^4-b^4)(r^2-s^2)\cos\epsilon\cos(2\phi+\epsilon)+2rsab(a^2-b^2)\sin(\gamma-\delta)\cos\epsilon\sin(2\phi+\epsilon) \\ & +\tfrac{1}{8}(a^2-b^2)^2\{r^2+s^2-2rs\cos(\gamma-\delta)\}\cos 2(2\phi+\epsilon).\end{aligned}$$

$$\begin{aligned}\frac{I}{c^2} = {} & \tfrac{1}{4}(r^2+s^2)\,\frac{(1+e^2)^2}{(1-e^2)^2}-2rs\,\frac{e^2}{(1-e^2)^2}\cos(\gamma-\delta)+\tfrac{1}{8}\{r^2+s^2+2rs\cos(\gamma-\delta)\}\cos 2\epsilon \\ & +\tfrac{1}{2}(r^2-s^2)\,\frac{1+e^2}{1-e^2}\cos\epsilon\cos(2\phi+\epsilon)+2rs\,\frac{e}{1-e^2}\sin(\gamma-\delta)\cos\epsilon\sin(2\phi+\epsilon) \\ & +\tfrac{1}{8}\{r^2+s^2-2rs\cos(\gamma-\delta)\}\cos 2(2\phi+\epsilon).\end{aligned} \qquad \ldots\ldots(12)$$

The intensity of illumination contains second and fourth harmonics; four maxima are therefore possible per revolution of the stage. The fourth harmonic I_4 is a function of the reflectivities and phase difference, the second harmonic I_2 is a function, in addition, of the ellipticity, and the relative setting of polarizer and analyser.

$$\frac{I_2}{c^2} = \tfrac{1}{2}\cos\epsilon\left\{(r^2-s^2)^2\,\frac{(1+e^2)^2}{(1-e^2)^2}+16r^2s^2\,\frac{e^2}{(1-e^2)^2}\sin^2(\gamma-\delta)\right\}^{1/2}$$

$$\frac{I_4}{c^2} = \tfrac{1}{8}\{r^2+s^2-2rs\cos(\gamma-\delta)\},$$

whence

$$\frac{I_2}{I_4} = 4\cos\epsilon\,\frac{\{(K^2-1)^2(1+e^2)^2/(1-e^2)^2+16K^2e^2/(1-e^2)^2\sin^2(\gamma-\delta)\}^{1/2}}{(K-1)^2+4K\sin^2\tfrac{1}{2}(\gamma-\delta)}. \qquad \ldots(13)$$

2.15. *Analyser and polarizer in crossed position.* When $\epsilon=\pi/2$ the second harmonic in the expression for the intensity vanishes and

$$\frac{I}{c^2}=\tfrac{1}{8}(r^2+s^2)\left\{2\frac{(1+e^2)^2}{(1-e^2)^2}-1\right\}-\tfrac{1}{4}rs\left\{1+\frac{8e^2}{(1-e^2)^2}\right\}\cos(\gamma-\delta)$$
$$-\tfrac{1}{8}\{r^2+s^2-2rs\cos(\gamma-\delta)\}\cos 4\phi.$$

This expression is a minimum for $\phi=0$, $\pi/2$, π, and $3\pi/2$ and a maximum for $\phi=\pi/4$, $3\pi/4$, $5\pi/4$ and $7\pi/4$. The minimum intensity is given by

$$\frac{I(0)}{c^2}=\frac{e^2}{(1-e^2)^2}\{r^2+s^2-2rs\cos(\gamma-\delta)\}.$$

Therefore when a given multi-grained specimen is rotated between crossed polarizers every grain will show four equal maxima and minima intensity per complete rotation.

2.16. *Analyser rotated from crossed position.* As the analyser is rotated from the crossed position the second harmonic term appears in the intensity and increases with the deviation from the crossed position. One pair of peaks is increased and the other pair decreased until, under certain conditions, the second pair of peaks disappears. These conditions are worked out below for the special case when the ellipticity of the principal waves is zero.

2.2. *Uniaxial Crystals*

2.21. *Uniaxial crystals.* For the general case of uniaxial crystals, those belonging to the tetragonal, hexagonal and trigonal systems, the ellipticity e is zero, i.e. the principal waves reduce to plane polarized waves in the two principal directions, corresponding to the given case of normal incidence.

The same is also true for special cases of non-uniaxial crystals; the normal to the surface must then lie in a certain conical surface, as explained in a previous publication (Mott and Haines 1951 a). To obtain the appropriate expressions for these cases e is put equal to zero in all the above results.

Thus the expression (9) for the intensity of the analysed reflected beam becomes $I_{\mathrm{R}}/c^2=\frac{1}{2}(r^2+s^2)+\frac{1}{2}(r^2-s^2)\cos 2\phi$, whence $I_{\mathrm{R}}(0)/I_{\mathrm{R}}(\pi/2)=r^2/s^2=K^2$. Equation (10) shows that the maximum and minimum values of I_{R} for uniaxial crystals occur along the principal directions of the section.

The variation of intensity with rotation of the analyser is obtained by putting $b=0$ in eqn. (11) and $e=0$ in (11 *a*):

$$I/c^2=\tfrac{1}{2}(r^2\cos^2\phi+s^2\sin^2\phi)+\tfrac{1}{2}\{(r^2\cos^2\phi-s^2\sin^2\phi)^2$$
$$+4r^2s^2\cos^2(\gamma-\delta)\cos^2\phi\sin^2\phi\}^{1/2}\cos 2(\psi-\psi_0) \qquad \ldots\ldots(14)$$

where the phase angle ψ_0 is given by

$$\tan 2\psi_0=\frac{2rs\cos\phi\sin\phi}{r^2\cos^2\phi-s^2\sin^2\phi}\cos(\gamma-\delta),$$

that is

$$\tan 2\psi_0=\frac{2K\tan\phi}{K^2-\tan^2\phi}\cos(\gamma-\delta). \qquad \ldots\ldots(14\,a)$$

The analyser gives the maximum illumination when $\psi=\psi_0$ and $\psi=\psi_0+\pi$. Minimum illumination occurs for analyser settings $\psi=\psi_0+\pi/2$ and $\psi=\psi_0+3\pi/2$. If the polarizer is set parallel to a principal axis $\phi=0$, and, therefore, $\psi_0=0$. For uniaxial crystals, therefore, maximum illumination occurs when the polarizers are parallel to a principal direction of the section.

The variation of intensity with the rotation of the stage is obtained by putting $e=0$ in eqns. (12) and (13):

$$I/c^2=\tfrac{1}{4}(r^2+s^2)+\tfrac{1}{8}\{r^2+s^2+2rs\cos(\gamma-\delta)\}\cos 2\epsilon+\tfrac{1}{2}(r^2-s^2)\cos\epsilon\cos(2\phi+\epsilon)$$
$$+\tfrac{1}{8}\{r^2+s^2-2rs\cos(\gamma-\delta)\}\cos 2(2\phi+\epsilon).$$

The intensity of illumination, as in the general case, contains both second and fourth harmonics; the ratio of the two harmonics is obtained by putting $e=0$ in (13):

$$\frac{I_2}{I_4}=4\cos\epsilon\,\frac{K^2-1}{(K-1)^2+4K\sin^2\tfrac{1}{2}(\gamma-\delta)}. \qquad \text{......(15)}$$

In the crossed position $\epsilon=\pi/2$; the intensity is then given by

$$I/c^2=\tfrac{1}{8}\{r^2+s^2-2rs\cos(\gamma-\delta)\}(1-\cos 4\phi).$$

The minimum intensity for uniaxial crystals with crossed polarizers is therefore zero.

2.22. *Analyser rotated from the crossed position.* Putting $2\phi+\epsilon=x$, the equation for the intensity becomes

$$I=I_0+I_2\cos x+I_4\cos 2x,$$

$$\frac{dI}{dx}=-(I_2+4I_4\cos x)\sin x, \qquad \frac{d^2I}{dx^2}=-I_2\cos x-4I_4(2\cos^2 x-1).$$

The turning points in the intensity occur when either (*a*) $\sin x_\mathrm{m}=0$ or (*b*) $\cos x_\mathrm{m}=-I_2/4I_4$. The first pair of turning points (*a*) always exist, the second pair (*b*) exist if $|I_2/4I_4|<1$; if $|I_2/4I_4|=1$, the two pairs coincide. The first pair of turning points occurs at $x_\mathrm{m}=0$ or π; $\cos x_\mathrm{m}=\pm 1$; $\cos 2x_\mathrm{m}=1$. Therefore $(d^2I/dx^2)_a=\mp I_2-4I_4=4I_4\{\pm|I_2/4I_4|-1\}$.

$I_4=\tfrac{1}{8}c^2\{r^2+s^2-2rs\cos(\gamma-\delta)\}>0$. Therefore $(d^2I/dx^2)_a$ is negative for both signs if $|I_2/4I_4|<1$, and is positive for one sign if $|I_2/4I_4|>1$. At the second pair of turning points

$$\left(\frac{d^2I}{dx^2}\right)_b=\frac{I_2^2}{4I_4}-4I_4\left(\frac{I_2^2}{8I_4^2}-1\right)=4I_4\left(1-\frac{I_2^2}{16I_4^2}\right)>0 \text{ if } \left|\frac{I_2}{4I_4}\right|<1,$$

that is, if the second pair exists. Hence if $|I_2/4I_4|<1$ there are two maxima in the range $0<x<2\pi$; if $|I_2/4I_4|>1$ there is only one maximum in the same range. As $x=2\phi+\epsilon$, the number of maxima in the range $0<\phi<2\pi$ are four and two respectively.

With eqn. (15), the condition for the existence of four maxima becomes

$$|\cos\epsilon|\,\frac{K^2-1}{(K-1)^2+4K\sin^2\tfrac{1}{2}(\gamma-\delta)}<1. \qquad \text{......(16)}$$

This condition will be satisfied for all values of ϵ if $K^2-1<(K-1)^2+4K\sin^2\tfrac{1}{2}(\gamma-\delta)$, i.e. if $2K\{1-2\sin^2\tfrac{1}{2}(\gamma-\delta)\}<2$ if $K\cos(\gamma-\delta)<1$. If, therefore, K, the ratio of the principal reflectivities, is less than $\sec(\gamma-\delta)$, four maxima will occur in the intensity–stage curve for all settings of the polarizers. If $K>\sec(\gamma-\delta)$, two of the maxima existing for crossed polarizers will decrease as $|\cos\epsilon|$ increases, and will disappear by coalescing with neighbouring minima at a setting obtained by treating inequality (16) as an equation. At this setting the intensity curve shows a very flat minimum, due to the vanishing of the first three derivatives.

An upper limit to the differential phase shift $\gamma-\delta$ can be obtained from eqn. (16). If the analyser setting $\frac{1}{2}\pi-\epsilon'$ just suppresses one pair of maxima, $\sin\epsilon'=\{K^2+1-2K\cos(\gamma-\delta)\}/(K^2-1)$ and

$$\cos(\gamma-\delta)=\tfrac{1}{2}(1-\sin\epsilon')K+(1+\sin\epsilon')/2K.$$

Considered as a function of K, $\cos(\gamma-\delta)$ is a minimum for

$$K^2=(1+\sin\epsilon')/(1-\sin\epsilon').$$

The minimum value is given by $\cos(\gamma-\delta)=(1+\sin\epsilon')/K=(1-\sin^2\epsilon')^{1/2}=\cos\epsilon'$. Hence the differential phase shift $\gamma-\delta$ cannot exceed the critical offset ϵ'.

2.23. *Angle through which the plane of polarization is rotated.* If the differential phase shift $\gamma-\delta$ in the case of a uniaxial crystal is small, the reflected beam is approximately plane polarized; this plane of polarization can be taken as coinciding with the analyser setting for maximum intensity as given by eqn. (14 *a*).

Assuming that $\gamma-\delta$ is small, this becomes $\tan 2\psi_0=2K\tan\phi/(K^2-\tan^2\phi)$, i.e. $\tan\psi_0=(\tan\phi)/K$. The rotation ω of the plane of polarization is $\phi-\psi_0$, so that $\tan\omega=(K-1)\tan\phi/(K+\tan^2\phi)$, which agrees with Capdecomme and Orcel (1941). The maximum value of this rotation occurs when $\tan^2\phi=K$, and the maximum rotation of the plane of polarization is given by $\tan\omega_{\max}=(K-1)/2\sqrt{K}$.

If $\gamma-\delta$ is small, more precisely if $\sin\frac{1}{2}(\gamma-\delta)\ll(K-1)/2\sqrt{K}$, the analyser setting at which one pair of intensity maxima disappear is given by

$$|\cos\epsilon|=(K-1)^2/(K^2-1)=(K-1)/(K+1).$$

If ϵ' is the departure from the crossed position $\epsilon=\frac{1}{2}\pi-\epsilon'$ then

$$\sin\epsilon'=(K-1)/(K+1).$$

Further if $K-1\ll 1$ $\quad \dfrac{\epsilon'}{\omega_{\max}}\simeq\dfrac{K-1}{K+1}\,\dfrac{2\sqrt{K}}{K-1}=\dfrac{2\sqrt{K}}{K+1}\simeq 1.$

Thus, provided that $\gamma-\delta\ll K-1\ll 1$, then, as the analyser is rotated from the crossed position, the intermediate maxima vanish for an offset equal to the maximum angle of rotation of the plane of polarization, a result found by Sampson (1929).

§ 3. EXPERIMENTAL OBSERVATIONS

3.1. *Procedure*

Specimens of tin, uranium, bismuth and stibnite were chosen for the tests and were prepared as follows:

Tin (*tetragonal*). Deformed by compression, mechanically polished and slowly cooled after recrystallization at 100° C to reduce the incidence of twinning. The surface was finally prepared by electrolytic polishing in 60% perchloric acid as recommended by Puttick (1949).

Uranium (*orthorhombic*). Cold-worked and recrystallized in the upper limit of the alpha-range. The surface was given a rough mechanical polish before electrolytic polishing in a solution containing 2 volumes sulphuric acid, 1 volume orthophosphoric acid and 2 volumes distilled water as described by Mott and Haines (1951 b).

Bismuth (*rhombohedral*). Deformed and recrystallized. Mechanically polished prior to 4 minutes electrolytic polishing in the solution described for uranium.

.B
5
2

Stibnite (*orthorhombic*). Polished mechanically on a lead lap and finished on soft felt using a suspension of rouge in oil.

The vertical illuminator on the microscope was a zinc sulphide coated glass slip and both the polarizer and analyser units were of polaroid sheet, the setting of the analyser being controlled and measured to an accuracy of about ±0·05°. A Wratten C filter (green) was used in conjunction with a mercury vapour lamp which was not stabilized against mains variations but showed only slight changes in intensity throughout a given experiment. The intensity of the reflected beam was measured by a photomultiplier tube of a microphotometer unit (AC Type 1175A constructed in the Electronics Division, A.E.R.E.) inserted in the monocular eyepiece tube of the microscope.

The microscope was set up by centring all objectives with respect to the rotating stage, which had been centred by the makers with respect to the optical axis. Using a highly polished, stainless steel specimen with the analyser, eyepiece and objective lens removed, the polarizer was rotated to give maximum intensity. This ensured that the plane of polarization was in the plane of incidence at the vertical reflector. The analyser was then rotated to give a new minimum with the lamp iris and field iris closed to minimum aperture, so that the specimen was illuminated by a narrow cone at near-normal incidence. The objective lens was then inserted and rotated in a special mount (which ensured that true centring was maintained for all settings) to a position in which minimum intensity was obtained. In this position the birefringence due to the strain in the objective was compensated as completely as possible by the stray polarization effects in the rest of the optical system, i.e. those due to the mounting plates used for the polaroids and the reflection at the glass slip illuminator. Under these conditions the four maximum intensities obtained on rotation of a single grain of an anisotropic metal under crossed polarizers should be equal. This was checked for each specimen and slight adjustments in the setting of the objective made, if necessary, to achieve this result.

A grain about 0·1–0·2 mm diameter and free from twinning was chosen in each specimen in turn. The iris diaphragm controlling the field aperture was closed down so that the image of the iris was totally enclosed within the chosen grain. In the first set of experiments intensity measurements were made as the stage was rotated through intervals of 10° except in the region of the maxima and minima intensities when some readings were taken every 5°. This was carried out for the crossed position and with the analyser rotated through various small angles up to a maximum of 4°. This gave experimental curves for the change in the intensity with azimuth angle ϕ for various values of ϵ' in the region of the crossed position. In addition, for some specimens, the position of the analyser was found at which the intensity was a minimum for various stage settings. From this the relationship between ω, the angle of rotation of the plane of polarization, and the azimuth angle ϕ was obtained.

3.2. *Results*

Tin. The curves shown in figs. 3 (*a*) and 3 (*b*) were obtained for two grains in the tin specimen differing markedly in orientation. From fig. 3 (*a*) it will be observed that for the crossed position the intensity–azimuth angle curve shows four peaks occurring at intensity values of 3·3, 3·4, 3·3 and 3·4 respectively. Alternate pairs of peaks are slightly lower than the others, which suggests that the system as a whole was not exactly equivalent to the crossed polarizer condition; this may have been due to a slight inaccuracy in setting the analyser or to incomplete compensation of the stray polarization effects in the microscope. The difference in alternative peaks in fig. 3 (*a*) is more marked in the curve obtained after the analyser has been rotated by 0·5° from the crossed position, the ratio between the pairs being about 2·5 : 1. After an analyser rotation of 1°, the intermediate peaks are reduced still further and are only just discernible. With further rotation of the analyser to 2° and 3° respectively from the crossed position, only

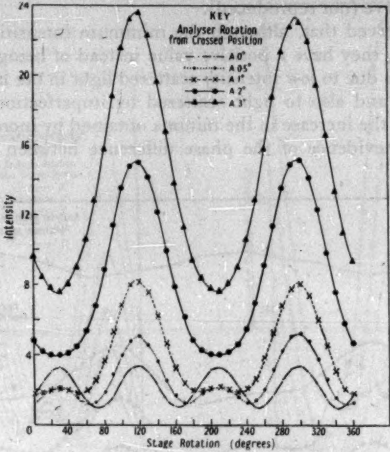


Fig. 3 (*a*). Electrolytically polished tin.

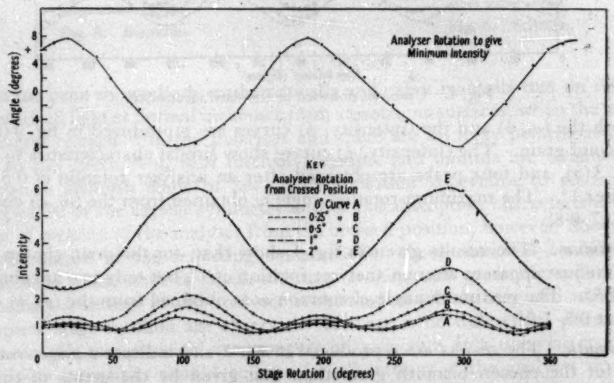


Fig. 3 (*b*). Tin electrolytically polished in perchloric acid solution.
Note. Upper scale should read −0·8 to 0·8 degrees, not −8 to 8.

two maxima are obtained and the minima have become more sharp, as predicted in §2.

Since the alternate maxima are almost eliminated at an analyser rotation of 1°, the probable maximum angle of rotation of the plane of polarization for this particular grain is of the order of 1·0°, which agrees with the value obtained from the (ω, ϕ) curve (not reproduced).

It will be observed that, although the minimum intensities for the crossed position are equal, they have a positive value instead of being zero as predicted by theory. This is due to low intensity scattered light in the laboratory reflected by the specimen, and also to light scattered by imperfections in the polished surface. Further, the increase in the minima obtained by increasing the rotation of the analyser is evidence of the phase difference between the two principal reflected beams.

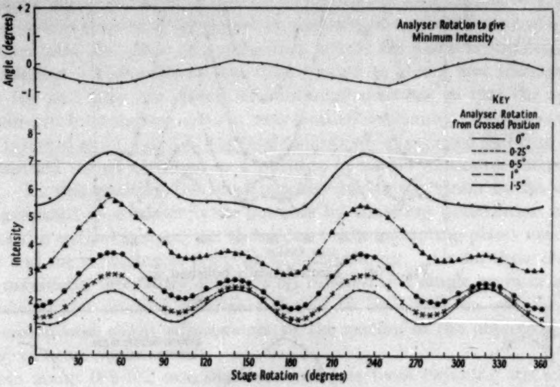


Fig. 4. Uranium polished in standard electrolyte.

Both the (ω, ϕ) and the (intensity, ϕ) curves are reproduced in fig. 3 (*b*) for the second grain. The (intensity, ϕ) curves show similar characteristics to those in fig. 3 (*a*), and four peaks are obtained after an analyser rotation of 0·5° and two after 1°. The maximum rotation angle ω obtained from the (ω, ϕ) curve is from 0·7–0·8°.

Uranium. The results given in fig. 4 show that, for the grain chosen, four peaks are just apparent after an analyser rotation of 1°, but only two are obtained after 1·5°. The maximum angle of rotation ω as obtained from the (ω, ϕ) curve is about 0·5–1·0°.

Bismuth. The results are reproduced in fig. 5 and indicate a larger rotation angle for the chosen bismuth grain than that given by the grains of tin and uranium. Four peaks were obtained after an analyser rotation of 3°, but two flat minima were evident after 4°. This agrees with the measured maximum angle of rotation ω of 3·6°.

Stibnite. The rotation angle for stibnite is usually greater than that for metals, and this is illustrated by the curves shown in fig. 6. Four maxima and minima were obtained after an analyser rotation of 4° from the crossed position but only two after a 6° rotation. The (ω, ϕ) curve is not reproduced but gave a maximum rotation angle ω for the grain investigated of about 5·2°. It is interesting to note that, as in the case of the metals, the minimum intensities increase slightly on rotation of the analyser, indicating different phase changes for the two principal components on reflection.

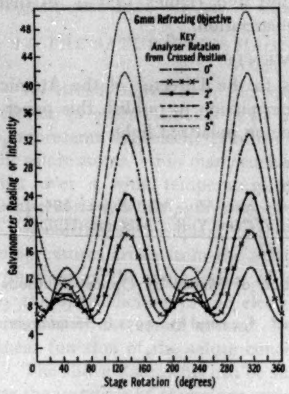


Fig. 5. Bismuth.

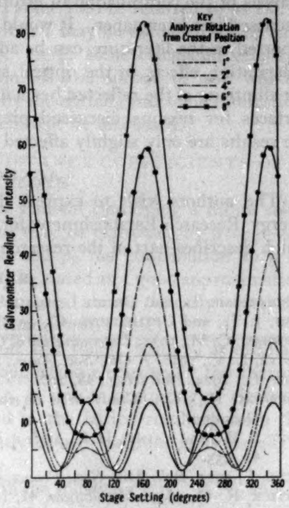


Fig. 6. Stibnite.

§ 4. DISCUSSION

It has been shown both mathematically and experimentally that on reflection of polarized light at normal incidence from anisotropic surfaces, when the analyser is exactly crossed with the polarizer and there are no stray polarization effects due to the instrument, four intensity maxima and minima are obtained for a change in azimuth angle of 2π. This condition appertains to all materials, irrespective of the crystal symmetry and of the absorption characteristics. The effect of rotation of the analyser from the crossed position, however, does depend on the nature of the reflecting surface. For those materials for which the differential phase shift is small, alternate maxima are gradually reduced until eventually a setting of the analyser is reached in which only two maxima per complete stage rotation are observed. If there is no differential phase change between the two principal components on reflection at a uniaxial crystal, the minimum intensities should be zero for all settings of the analyser up to that which just eliminates the alternate maxima; if the differential phase change is small, the minimum intensities will increase slightly with offset. If the differential phase shift is sufficiently large, then four maxima and minima will be obtained for all settings of the analyser.

The above results apply only to optical systems which are free from stray polarization effects, and, in general, ideal conditions are not obtained. The main sources of uneven polarization over the field of view are strain in the objective and polarizer mounts and refraction and reflection at oblique incidence to the vertical reflector unit. In the present work the setting of the objective was such that the elliptical effects due to strain partially corrected for the non-uniform polarization resulting from the slip illuminator. The precise effect on the intensity analysis of the combination of errors due to defects in the optical system will be discussed in a later paper. It would appear, however, that many of the anomalies reported in the literature can be adequately explained on the basis of the stray polarization effects in the optical systems used. In the case of etched metals the ellipticity of the reflected beam is considerably more marked than for polished surfaces for reasons discussed previously (Mott and Haines 1951 a), so that the results are only slightly affected by stray polarization.

ACKNOWLEDGMENTS

The authors wish to express their thanks to the Director of the Atomic Energy Research Establishment for his kind permission to publish this paper, which describes part of the research programme on polarized light.

REFERENCES

Capdecomme, L., and Orcel, J., 1941, *L'Optique*, **20,** 47.
Eash, J. T., and Upthegrove, C., 1933, *Trans. Amer. Inst. Min. Met. Engrs.*, **104,** 221.
Farnham, C. M., 1931, *Determination of Opaque Minerals* (New York : McGraw-Hill).
Foster, L. V., 1938, *J. Opt. Soc. Amer.*, **28,** 124 and 127.
Jones, O., 1924, *Phil. Mag.*, **48,** 207.
Morrogh, H., 1949, Conference on the Examination of Metals by Optical Methods (London).
Mott, B. W., and Haines, H. R., 1951 a, *Research, Lond.*, **4,** 24 and 63; 1951 b, *Metallurgia*, **43,** 255.
Orcel, J., 1925, *Bull. Min. Soc., France*, **48,** 272.
Puttick, K. E., 1949, *Metallurgia*, **41,** 120.
Sampson, E., 1923, *Econom. Geol.*, **18,** 775; 1929, *Ibid.*, **24,** 412.
Short, N. M., 1940, U.S. Department of the Interior Geological Survey Bulletin No. 914.
Szivessy, G., 1929, *Handbuch Phys.*, **20,** 861.
Wright, F. E., 1919, *Proc. Amer. Phil. Soc.*, **58,** 401.

The Magneto-Resistance of Ferromagnetic Al–Si–Fe Alloys

By R. PARKER
Department of Physics, University of Nottingham

Communicated by L. F. Bates ; MS. received 7th April 1952

ABSTRACT. By extending the treatment in an earlier paper an equation is derived for the saturation magneto-resistance of some mixed ferromagnetic alloys as a function of composition and temperature. Results calculated from this equation are in good agreement with experiment for the Al–Si–Fe alloy system.

§ 1. INTRODUCTION

In a recent publication (Parker 1951 a, referred to as I) the author has suggested that by making a number of plausible assumptions an equation can be derived for the saturation magneto-resistance coefficient $(\Delta\rho/\rho)_{\text{sat}}$ of certain ferromagnetic binary alloys; $(\Delta\rho/\rho)_{\text{sat}}$ may be defined as the relative change in electrical resistivity when an initially demagnetized material is magnetized to technical saturation. The equation derived describes $(\Delta\rho/\rho)_{\text{sat}}$ as a function of temperature and composition of the alloy. There is good quantitative agreement between the calculated value of $(\Delta\rho/\rho)_{\parallel\,\text{sat}}$, which is the value of $(\Delta\rho/\rho)_{\text{sat}}$ when the magnetic field is applied along the direction of current flow in a specimen, and the experimental values for alloys of iron with silicon (Shirakawa 1939) and iron with aluminium (Parker 1951 b, referred to as II).

It is shown in § 2 below that the treatment in I may be extended to the case of alloys containing more than two elements. In § 3 some calculated values of $(\Delta\rho/\rho)_{\parallel\,\text{sat}}$ for dilute solid solutions of aluminium and silicon in iron will be compared with experimental values obtained by the author, and some interesting magneto-resistance features of this alloy system will be discussed.

§ 2. THE SATURATION MAGNETO-RESISTANCE COEFFICIENTS OF SOME MIXED ALLOYS

As already noted in I, the resistivity ρ of a metal containing a dilute solid solution of other elements in its crystal lattice may be expressed by $\rho=\rho_{\mathrm{F}}+\rho_{\mathrm{S}}$; ρ_{F} represents the resistivity of the pure metal and ρ_{S} the contribution to ρ of the solute atoms. For many such alloys it may be stated as a good approximation that over a wide temperature range ρ_{F} increases almost linearly with the temperature t, while ρ_{S} is independent of t (Matthiessen's rule).

According to the quantum theory (cf. Mott and Jones 1936), ρ_{F} at ordinary temperatures arises from the scattering of conduction electrons by the thermal vibrations of the ionic lattice of the metal; ρ_{S} is due to scattering of electrons by the aperiodicities in the electrostatic field in the lattice resulting from the presence of the solute atoms. For sufficiently dilute solutions ρ_{S} is usually a linear function of the solute concentration.

The above statements can be summarized in the following way. If N_{F} and N_i are the numbers of atoms per cm³ of solvent metal and of the ith solute element respectively in the lattice, then

$$\rho=N_{\mathrm{F}}\sigma_{\mathrm{F}}+\Sigma_i N_i\sigma_i \quad \text{and} \quad N_{\mathrm{F}}\gg\Sigma_i N_i, \qquad \ldots\ldots(1)$$

where σ_{F} and σ_i are measures of the efficiency of scattering of conduction electrons, averaged over all crystal directions, due to thermal vibrations of the lattice and to solute atoms of the ith element respectively. As in I, the contribution to ρ of the thermal vibrations of the solute atoms is neglected. Let Δ again denote the changes when a specimen is brought from the demagnetized state to technical saturation; then, by definition,

$$(\Delta\rho/\rho)_{\text{sat}}=(N_{\mathrm{F}}\Delta\sigma_{\mathrm{F}}+\Sigma_i N_i\Delta\sigma_i)/(N_{\mathrm{F}}\sigma_{\mathrm{F}}+\Sigma_i N_i\sigma_i).$$

This equation reduces to

$$(\Delta\rho/\rho)_{\text{sat}}=[\rho_{\mathrm{F}}(\Delta\rho_{\mathrm{F}}/\rho_{\mathrm{F}})+\Sigma_i\rho_i(\Delta\rho_i/\rho_i)]/[\rho_{\mathrm{F}}+\rho_{\mathrm{S}}].$$

The last equation can now be applied to the case of dilute solutions of aluminium and silicon in iron. Let the suffixes 1 and 2 refer to the elements aluminium and silicon respectively, then

$$(\Delta\rho/\rho)_{\text{sat}}=[\rho_{\mathrm{F}}(\Delta\rho_{\mathrm{F}}/\rho_{\mathrm{F}})+\rho_1(\Delta\rho_1/\rho_1)+\rho_2(\Delta\rho_2/\rho_2)]/(\rho_{\mathrm{F}}+\rho_{\mathrm{S}}). \qquad \ldots\ldots(2)$$

For the case of the longitudinal saturation magneto-resistance all quantities on the right-hand side of eqn. (2) can be obtained from the results of independent experiments, and contain no unknown parameters. Thus direct comparison betwen $(\Delta\rho/\rho)_{\parallel\,\text{sat}}$ obtained from eqn. (2) and the experimental value can be made.

In the calculated curves of § 3 below, the value of ρ_F is the mean quoted for iron by Cleaves and Thompson (1935), and $\Delta\rho_F/\rho_F$ is obtained from the experimental results of Shirakawa (1939, 1940); $\Delta\rho_1/\rho_1$ and $\Delta\rho_2/\rho_2$ are taken from the results on the binary alloys reported in I and II. Values of ρ_S were obtained from resistivity measurements, and it was ascertained that ρ_S of each alloy considered is approximately independent of temperature. The only novel assumption here arises from the estimate of ρ_1 and ρ_2. According to eqn. (1) solutions of two separate elements contribute independently to ρ_S. If this equation applies to the aluminium–silicon–iron system it is possible to write

$$\rho_S = \rho_1 + \rho_2. \qquad \text{......(3)}$$

The results of a very extensive investigation at room temperature by Masumoto (1936) suggest that eqn. (3) is closely obeyed for solute concentrations not exceeding 7% by weight. Resistivity measurements on the specimens discussed in § 3 are in good agreement with Masumoto's results. The values of ρ_1 and ρ_2 in eqn. (2) have therefore in each case been taken from the well-known results for binary alloys.

§ 3. EXPERIMENTAL RESULTS AND DISCUSSION

Measurements of longitudinal magneto-resistance in fields of up to 1 200 oersteds and in the temperature range −78°c to +300°c were made on three alloys. The chemical analysis, in per cent by weight, of the alloys is given in the table:

Alloy	Al	Si	C	P	Mn	S	Cu	Ni
A	0·45	0·46	0·01	0·02	0·12	0·02	0·09	0·06
B	0·44	1·24	0·01	0·02	0·16	0·01	0·08	0·05
C	0·96	0·81	0·01	0·03	0·13	0·02	0·08	0·04

Values of $(\Delta\rho/\rho)_{\parallel\,\mathrm{sat}}$ were estimated in the usual manner. The results, together with the calculated curves, are shown in fig. 1. The agreement obtained is about the same as that reported in I and II and is well within the limits of the experimental errors in the determination of the quantities used in (2).

The transverse saturation magneto-resistance coefficient $(\Delta\rho/\rho)_{\perp\,\mathrm{sat}}$ of the alloys for magnetization perpendicular to the direction of current flow was also measured between 15°c and 300°c, the results being recorded in fig. 2. It was pointed out in II that direct comparison with eqn. (2) is not possible owing to lack of data of $(\Delta\rho/\rho)_{\perp\,\mathrm{sat}}$ for pure iron. A short calculation, however, shows that (analogous to the results for aluminium–iron alloys) the variation of $(\Delta\rho/\rho)_{\perp\,\mathrm{sat}}$ in the above temperature range should not exceed $0{\cdot}3 \times 10^{-3}$ for any of the above alloys. It can be seen from fig. 2 that this is the case.

So far discussion has been confined to the case of magneto-resistance coefficients associated with technical saturation. Unusual features in the longitudinal magneto-resistance coefficient $(\Delta\rho/\rho)_{\parallel}$ below technical saturation were observed in the cases of alloys B and C. For most ferromagnetic materials $(\Delta\rho/\rho)_{\parallel}$ increases or decreases steadily with H until saturation is reached if the usual experimental errors are eliminated (Stierstadt 1930). For the above two alloys, however, it was found that the $[(\Delta\rho/\rho)_{\parallel}, H]$ curves exhibit a marked initial minimum, which becomes less pronounced with increasing temperature. Typical curves are shown in fig. 3. Somewhat similar characteristics have previously been reported by Shirakawa (1939) for 1·66% Si–Fe, and by Parker (1950) for 1·8% Si–Fe and for a single crystal of 2·8% Si–Fe.

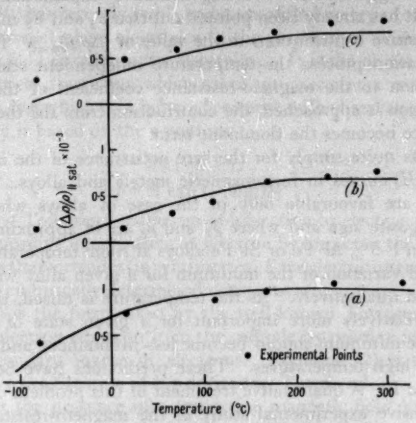


Fig. 1. $(\Delta\rho/\rho)_{\parallel\,\mathrm{sat}}$ plotted against temperature.

(*a*) Alloy A: (*b*) Alloy B: (*c*) Alloy C.

Curves are calculated from eqn. (2).

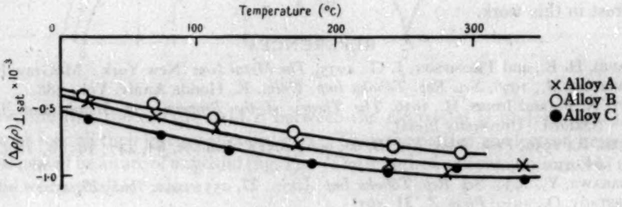


Fig. 2. $(\Delta\rho/\rho)_{\perp\,\mathrm{sat}}$ plotted against temperature.

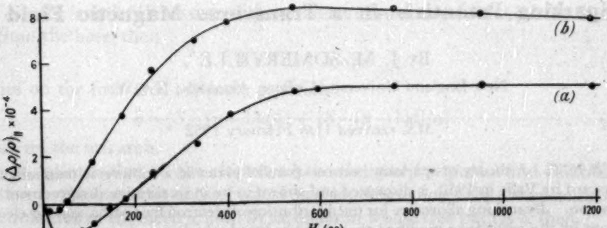


Fig. 3. The longitudinal magneto-resistance of 0·96% Al–Fe, 0·81% Si–Fe.

(*a*) 18°c, (*b*) 242°c.

The writer suggests the following tentative explanation of the above phenomenon. It has already been pointed out that σ_S and σ_F make respectively negative and positive contributions to the value of $(\Delta\rho/\rho)_{\parallel\,\mathrm{sat}}$. In the early part of the magnetization process the temperature independent scattering makes a larger contribution to the magneto-resistance coefficient of the alloy, but, as technical saturation is approached, the contribution from the thermal vibrations of the iron lattice becomes the dominant term.

This accounts quite simply for the rare occurrence of the initial minimum in the $[(\Delta\rho/\rho)_{\parallel}, H]$ curves in ferromagnetic metals and alloys. Conditions for this occurrence are favourable only in the case of alloys where $\Delta\rho_F/\rho_F$ and $\Delta\rho_S/\rho_S$ are of opposite sign and where ρ_F and ρ_S are of approximately the same magnitude, i.e. in 1–3% Al–Fe or Si–Fe alloys at room temperature.

The observed variation of the minimum for a given alloy with temperature may be predicted qualitatively. As the temperature is raised, the contribution of ρ_F becomes relatively more important for a given state of magnetization. Consequently the minimum should become less pronounced and occur in lower applied fields at high temperatures. These predictions have been verified for both alloys A and B. A quantitative treatment of this problem cannot be given without an extensive experimental study of the magneto-resistance properties and the magneto-crystalline energy of single crystals as a function of temperature and composition.

ACKNOWLEDGMENTS

The author thanks Professor L. F. Bates and Dr. R. Street for their interest in this work.

REFERENCES

Cleaves, H. E., and Thompson, J. G., 1935, *The Metal Iron* (New York : McGraw-Hill).

Masumoto, H., 1936, *Sci. Rep. Tôhoku Imp. Univ.*, K. Honda Anniv. Vol., 388.

Mott, N. F., and Jones, H., 1936, *The Theory of the Properties of Metals and Alloys* (Oxford : University Press).

Parker, R., 1950, *Proc. Phys. Soc.* B, **63**, 996; 1951a, *Ibid.*, A, **64**, 447; 1951b, *Ibid.*, B, **64**, 930.

Shirakawa, Y., 1939, *Sci. Rep. Tôhoku Imp. Univ.*, **27**, 255; 1940, *Ibid.*, **29**, 152.

Stierstadt, O., 1930, *Phys. Z.*, **31**, 561.

Sparking Potentials in a Transverse Magnetic Field

By J. M. SOMERVILLE

New England University College, Armidale, N.S.W.

MS. received 11*th February* 1952

ABSTRACT. A theory of sparking between parallel plates in a transverse magnetic field, put forward by Valle in 1950, is discussed and shown to be in qualitative disagreement with observation. By making allowance for the distribution of electron free paths, and for electron recapture by the cathode, a modified theory is produced which is in better agreement with observation. Further refinement of the theory requires an approximate expression for the first Townsend coefficient which is valid when a magnetic field is present. The possibility of obtaining such an expression is discussed.

§1. INTRODUCTION

Valle (1950) has recently published a simplified theory of sparking between infinite parallel plane electrodes in the presence of a uniform transverse magnetic field parallel to the electrodes. It is the object of this paper to outline a theoretical treatment which takes into account a number of factors ignored by Valle and which is in better agreement with observation.

Valle's theory is based on the equations

$$\alpha d = \log(1 + 1/\gamma) \qquad \ldots\ldots(1)$$

and

$$\alpha/p = A\exp(-Bp/E) \qquad \ldots\ldots(2)$$

where α is the first Townsend coefficient of ionization by electron collision, γ is the probability of a positive ion releasing an electron by impact at the cathode, d is the distance between the electrodes (in centimetres), E the electric intensity (in volts/cm) and A and B are empirically determined constants.

Equation (1) is the form taken by the well-known Townsend condition for sparking when electron emission from the cathode by positive ion bombardment is the chief secondary source of electrons. Equation (2) is the approximate expression for α/p as a function of E/p derived by Townsend.

To allow for the presence of a transverse magnetic field Valle replaces the pressure p in eqn. (2) by an 'equivalent pressure' p_e derived as follows.

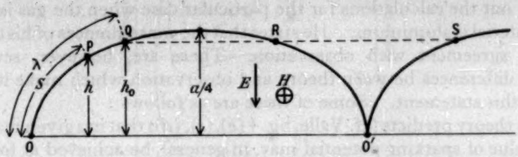


Fig. 1.

Assume that the electric field E between the electrodes is uniform and that electrons lose all their energy at collision. Then the path of an electron between collisions will be an arc of a cycloid (fig. 1). The length of one complete arch of this cycloid we may call a,[*] where

$$a = 8 \times 10^8 E\boldsymbol{m}/\boldsymbol{e}H^2, \qquad \ldots\ldots(3)$$

H being the magnetic intensity in oersteds and $\boldsymbol{e}$ and $\boldsymbol{m}$ the charge (e.m.u.) and the mass (g) of an electron.

If S is the length OP measured along the arc of such a cycloid and h the distance of P from the base, then

$$h = S(1 - S/a) \qquad \ldots\ldots(4)$$

if P lies on the first arch or, more generally,

$$h = \{S - (n-1)a\}[1 - \{S - (n-1)a\}/a] \qquad \ldots\ldots(5)$$

if P lies on the nth arch.

Valle assumes that all the electrons have *the same free path* λ and confines his attention to the first arch of the cycloid so that, if h_0 denotes the distance travelled in the direction of the electric field by an electron whose free path is λ, then

$$h_0 = \lambda(1 - \lambda/a). \qquad \ldots\ldots(6)$$

---

[*] Valle denotes the length of a complete arch of the cycloid by $8a$. Throughout this paper Valle's results are expressed in the notation most suitable to our purpose.

He then takes as the 'equivalent pressure'

$$p_e = p\lambda/h_0 \qquad \text{......(7)}$$

$$= L/h_0, \qquad \text{......(8)}$$

where

$$L = p\lambda, \qquad \text{......(9)}$$

and is thus equal in magnitude to the mean free path at unit pressure. The equivalent pressure p_e is inserted in eqn. (2) instead of the actual pressure p, giving

$$\alpha h_0 = AL\exp(-BL/Eh_0). \qquad \text{......(10)}$$

If V is the potential difference between the electrodes expressed in volts then $E = V/d$ and (10) may be written

$$\alpha h_0 = AL\exp(-BLd/Vh_0). \qquad \text{......(11)}$$

Elimination of α between (1) and (11) gives

$$\frac{h_0}{d} = \frac{AL\exp(-BLd/Vh_0)}{\log(1+1/\gamma)} \qquad \text{......(12)}$$

For a given gas and cathode material, i.e. for given A, B, L and γ, eqns. (3), (6), (9) and (12) determine the sparking potential V as a function of pd and Hd. Valle has devised a convenient scheme for the computation of V from these equations and has carried out the calculations for the particular case when the gas is air and the cathode material is aluminium. He states that the consequences of his theory are in qualitative agreement with observation. There are, however, several major qualitative differences between theory and observation which make it difficult to agree with this statement. Some of these are as follows :

(i) The theory predicts (cf. Valle, fig. 4 (*b*), (*c*), (*d*)) that in a given magnetic field the same value of sparking potential may, in general, be achieved at four different pressures. There does not appear to be any observational support for this, and observers like Meyer (1919), Wehrli (1922) and Penning (1937) agree in finding, for a given magnetic field, only two pressures corresponding to each value of the sparking potential.

(ii) The theory predicts that under some circumstances (cf. Valle, fig. 4 (*b*)) a range of pressures exists for which no sparking at all is possible at any voltage, although sparking may take place at pressures above or below this range. This again lacks observational support.

(iii) At low values of the product *pd* the theory yields curves of sparking potential against magnetic field strength (cf. Valle, fig. 4 (*c*)) which move to the left, i.e. towards lower magnetic fields, as *pd* decreases, and which in general intersect one another or else lie entirely outside one another. Figure 2 is a sketch showing the type of (sparking potential, magnetic field strength) curves obtained experimentally by Penning (1937) in argon between coaxial cylindrical electrodes. It is seen that the curves corresponding to different pressures do not intersect and that any curve corresponding to a lower pressure lies entirely within any curve corresponding to a higher pressure. Moreover, the curves move to the *right* as the pressure decreases.

(iv) The theory predicts that, no matter how low the pressure may be, the minimum sparking potential for a given pressure may always be reduced to the same value by a suitable choice of magnetic field. Figure 2 shows that the observed minimum attainable sparking potential increases rapidly as the pressure is lowered.

We now seek the physical reasons for these points of disagreement between theory and observation with a view to the development of a theory more in line with experiment. In so doing it will be convenient to make one modification at a time and follow out its implications rather than to make several modifications at once.

§ 2. THE DISTRIBUTION OF FREE PATHS

The first and second points of disagreement between Valle's theory and experiment are directly attributable to neglect of the distribution of free paths, the assumption being made that all electrons have the same free path between collisions. This may be seen as follows :

For any given value of V, eqn. (12) in general determines two values of h_0. Each of these values of h_0 determines, through eqn. (6), two values of λ (the arcs OQ and OR in fig. 1) and hence two values of p, giving, in all, four* values of p corresponding to each value of V.

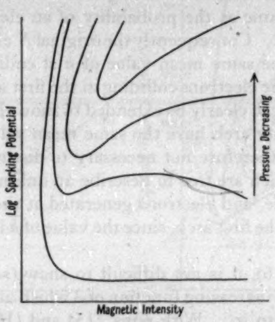


Fig. 2. Sketch (not to scale) showing types of (sparking potential, magnetic intensity) curves for different gas pressures obtained by Penning (1937) in argon between coaxial cylindrical electrodes in an axial magnetic field.

Suppose now that we recognize that all electrons do not have the same free path, but that the probability of a free path length between x and $x+dx$ is $\exp(-x/\lambda)dx/\lambda$. We would then replace h_0 by $\bar{h}$, the mean value of h at collision, giving, instead of (8),

$$p_0 = L/\bar{h}. \qquad \ldots\ldots(13)$$

Equations (10) and (12) then become, respectively,

$$\alpha\bar{h} = AL\exp(-BL/E\bar{h}) \qquad \ldots\ldots(14)$$

$$\frac{\bar{h}}{d} = \frac{AL\exp(-BLd/V\bar{h})}{\log(1+1/\gamma)}. \qquad \ldots\ldots(15)$$

The value of $\bar{h}$ in terms of λ and a has been given by Kugler, Ollendorf and Roggendorf (1933), and we may write it in the form

$$\bar{h} = \lambda\{\coth(a/2\lambda) - 2\lambda/a\}, \qquad \ldots\ldots(16)$$

which is the equivalent equation replacing (6).

---

* In using an equation equivalent to (6) Valle has arbitrarily limited consideration to the first cycloidal arch, but in general if an electron completes such an arch without collision it will immediately begin another identical one, and other values of λ (e.g. the arc OPQRO′S) corresponding to the value h_0 become possible, leading to an infinite number of values of the pressure corresponding to each value of V.

This equation may also be conveniently written in the form

$$\bar{h}/a=(\lambda/a)\{\coth(a/2\lambda)-2\lambda/a\} \qquad \ldots\ldots(16\,a)$$

or

$$\bar{h}/a=F(\lambda/a) \qquad \ldots\ldots(16\,b)$$

where $F(x)=x\{\coth(1/2x)-2x\}$, giving $\bar{h}/a$ as a function of λ/a. Equation (16) was derived on the assumption that the electrons were free to go on describing an indefinite number of cycloidal arches until collision. It is easy to see, however, that the mean value of h for the electrons which collide in the first arch is the same as the mean value for the whole group. Consider a group of N electrons starting from rest at O (fig. 1). This group may be divided into two subgroups, namely the $N\exp(-a/\lambda)$ electrons which reach O′ without colliding, and the $N\{1-\exp(-a/\lambda)\}$ electrons which collide in the first arch. Now the $N\exp(-a/\lambda)$ electrons reaching O′, and beginning a new arch from there, differ only in total number from the original N starting from O; their chance per unit distance of colliding is the same, and the probability of an electron starting from O′ reaching a certain distance before collision is the same as the probability of an electron reaching the same distance starting from O. Consequently the original N electrons and the subgroup of $N\exp(-a/\lambda)$ have the same mean value of h at collision. Therefore the remaining subgroup, i.e. the electrons colliding in the first arch, have this same mean value. (This argument can clearly be extended to show that the electrons colliding in the second or any other arch have the same mean value of h at collision.) In applying eqn. (16) it is therefore not necessary to distinguish between electrons generated in the gas, which are free to describe an unlimited number of cycloidal arches before they collide, and electrons generated at the cathode, which may be recaptured at the end of the first arch, since the value of $\bar{h}$ is the same for either class of electrons.

Returning to eqn. (16), it is not difficult to show (see Appendix) that $\bar{h}$ is a single-valued monotonic increasing function of λ which asymptotically approaches the limit $a/6$ as λ tends to ∞. With eqns. (15) and (16) replacing (12) and (6) respectively there will be only two values of the pressure corresponding to each value of V, because (15) will determine two values of $\bar{h}$ for each of V, and (16) will determine only one value of λ, and hence of p, for each value of $\bar{h}$. This modification therefore removes our first objection to Valle's theory.

It removes the second also, as we shall now see. In eqn. (15) the exponential cannot exceed unity, and hence no sparking is possible if $\bar{h}>ALd/\log(1+1/\gamma)=h_{\mathrm{m}}$ say. Since $\bar{h}$ is a monotonic function of λ, and hence of p, once p is so low that $\bar{h}>h_{\mathrm{m}}$, further reduction in p will only result in further increase in $\bar{h}$, so that, once sparking becomes impossible, it remains impossible at lower pressures.

Some (V, Hd) curves for various values of pd computed from eqns. (15) and (16) are shown in fig. 3. In computing these and other sparking potential curves in this paper the following values of the constants have been used, appropriate to air as the gas and copper as the cathode material, namely $A=14{\cdot}6\,(\mathrm{cm\,mm\,Hg})^{-1}$, $B=365$ v/cm mm Hg, as given by Townsend (1915), $L=0{\cdot}026$ cm mm Hg, $\gamma=0{\cdot}025$, as given by Engel and Steenbeck (1932). It will be seen that our third objection has also been partially removed in that the curves no longer move to the left with decreasing pressure, but the fourth objection still remains, and the modified theory still implies that the sparking potential can be reduced to the same minimum value for all values of pd. Moreover the sparking potential curves now tend to a limiting form as the pressure is reduced, and there is little change in the curves for values of *pd* below about 0·01 cm mm Hg. This independence of pressure at low pressures is due to the fact, already discussed, that $\bar{h}$ tends to a limiting value $(a/6)$ as λ increases.

Physically this independence of pressure is due to the assumption that, in the magnetic field, electrons keep on describing a series of cycloidal arches until they collide with gas atoms. If the pressure is lowered, electrons describe more arches before making a collision, but the collision is ultimately made, however low the pressure may be, and is just as effective in promoting sparking as if it had been made earlier. The situation would be changed if a process existed whereby some electrons were removed from the space between the electrodes before collision and if the chance of such removal increased as the pressure decreased. One such process of removal is by electron recapture at the cathode.

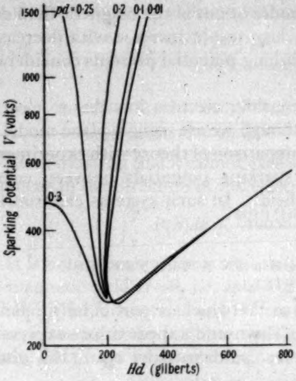


Fig. 3. Calculated sparking potential V for sparking between parallel plane copper electrodes a distance d cm apart in air at a pressure p mm Hg in the presence of a transverse magnetic field H oersted, allowance being made for the distribution of electronic free paths.

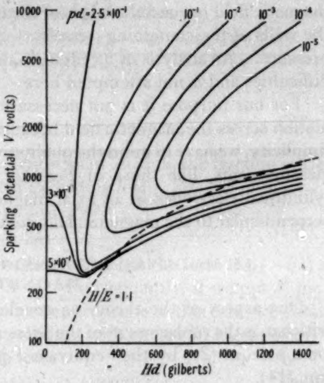


Fig. 4. Calculated sparking potentials, for the same conditions as those of fig. 3, when allowance is made both for the distribution of electronic free paths and for electron recapture by the cathode.

At points on the dotted line the condition $H/E\ (=Hd/V)=1{\cdot}1$ is satisfied. Its significance is discussed in the text.

§ 3. ELECTRON RECAPTURE AT THE CATHODE

Electrons generated at the cathode which describe a complete cycloidal arch without making a collision will in general be recaptured by the cathode at the end of the arch. Thus, of N electrons leaving the cathode, $N\exp(-a/\lambda)$ will complete an arch of length a without collision and be recaptured by the cathode, and, so far as their influence on breakdown is concerned, they might as well never have existed. Consequently, of every N electrons generated at the cathode only $N\{1-\exp(-a/\lambda)\}$ are effective in promoting breakdown, so that the coefficient of emission of electrons by positive ions at the cathode is effectively reduced by the factor $1-\exp(-a/\lambda)$. We can therefore make allowance for electron recapture at the cathode by substituting $\gamma\{1-\exp(-a/\lambda)\}$ for γ in eqn. (15), obtaining

$$\frac{\bar{h}}{d}=\frac{AL\exp(-BLd/V\bar{h})}{\log[1+1/\gamma\{1-\exp(-a/\lambda)\}]}. \qquad \ldots\ldots(17)$$

This equation, together with eqns. (3) (9) and (16), now gives values of the sparking potential with allowance both for the distribution of electron free paths and electron recapture at the cathode.

Figure 4 shows curves of sparking potential plotted against Hd derived from eqn. (17). It will be noted that in these curves the desired progression of the minima upward and from left to right with decreasing pressure has been achieved, and that the qualitative agreement with Penning's curves (fig. 2) is better.

§ 4. OTHER SOURCES OF ELECTRON LOSS

Recapture by the cathode is not the only process causing loss of electrons from the space between the electrodes. Any actual discharge tube with plane parallel electrodes must be finite, and so must be the extent of the magnetic field. Electrons may be lost as effective agents in promoting sparking if they diffuse along the magnetic field beyond the edges of the electrodes or out of the magnetic field or to the walls of the containing vessels. Such loss would increase with decreasing pressure. An analysis of its effect on the sparking potential presents considerable difficulties and is not attempted here.

For our purpose it is not necessary to consider electron loss due to electron motion across the magnetic field because, although we are using a plane model for simplicity, we have in mind the qualitative comparison of theory with experimental measurements, like those of Penning, of sparking potentials between coaxial cylindrical electrodes in an axial magnetic field. In such systems electron loss perpendicular to the magnetic field does not occur.

§ 5. MODIFIED EXPRESSION FOR α IN A TRANSVERSE MAGNETIC FIELD

One aspect of the theory as developed so far which is particularly open to criticism is the replacement of the pressure in Townsend's approximate expression for α/p (eqn. (2)) by the 'equivalent pressure' p_e defined by eqn. (13), giving eqn. (14).

Equation (2) is derived as follows. Assume that electrons lose all energy at collision, that an electron with energy greater than eW' ionizes at collision and that one with less does not. In one centimetre path length an electron will make on the average $1/\lambda'$ collisions. (The symbol λ' is used for the mean free path in recognition of the fact that values of λ' derived from eqn. (2) are not equal to the kinetic theory mean free path or the mean free path measured in other ways, and W' is used for similar reasons). Of these $1/\lambda'$ collisions a fraction $\exp(-l/\lambda')$ will be preceded by a free path exceeding l. Assuming the electric field to be uniform and equal to E, then if we put $l = W'/E$, we have the number of ionizing collisions per centimetre, namely

$$\alpha = (1/\lambda') \exp(-W'/E\lambda') \qquad \ldots\ldots(18)$$

or

$$\alpha/p = (1/L') \exp(-W'p/EL'), \qquad \ldots\ldots(19)$$

writing L' for $p\lambda'$. This leads to eqn. (2) if we write

$$A = 1/L', \quad B = W'/L'. \qquad \ldots\ldots(20)$$

If we try to follow the same line of argument when a transverse magnetic field is present, we find that we cannot simply replace the free path l by the free distance h travelled in the direction of the electric field. This is so because values of l and h are not distributed in the same way, and it is not true to say that the fraction of free distances in the direction of the electric field is equal to $\exp(-h/\bar{h})$.

As a preliminary to obtaining an equivalent equation to (2) when a magnetic field is present, we will now derive the distribution function for h. It is clear, from what has already been said, that we need only calculate the distribution function for electrons colliding in the first arch, since the distribution function for electrons colliding in other arches will be the same. Referring to fig. 1, an electron starting from rest at O will be at a given distance h from the base line twice, at distances from O which we may call S_1 and S_2, and which are the roots of eqn. (4) :

$$S_1 = \tfrac{1}{2}a\{1-(1-4h/a)^{1/2}\}, \qquad S_2 = \tfrac{1}{2}a\{1+(1-4h/a)^{1/2}\}. \qquad \ldots\ldots(21)$$

Of N electrons starting from O, the number colliding at a distance from the base line between h and $h+dh$ is $N\exp(-S_1/\lambda)\,|dS_1|/\lambda + N\exp(-S_2/\lambda)\,|dS_2|/\lambda$, where dS_1 and dS_2 correspond to dh, and, from (21), are equal to $\pm dh/(1-4h/a)^{1/2}$.

The total number of electrons colliding in the first arch is $N\{1-\exp(-a/\lambda)\}$, so that the probability of collision within the interval h to $h+dh$ is

$$P(h)dh = \frac{[N dh \exp(-a/2\lambda)\,[\exp\{a(1-4h/a)^{1/2}/2\lambda\} + \exp\{-a(1-4h/a)^{1/2}/2\lambda\}]}{\lambda(1-4h/a)^{1/2}N\{1-\exp(-a/\lambda)\}} \qquad \ldots\ldots(22)$$

$$= \frac{\cosh\{(a/2\lambda)(1-4h/a)^{1/2}\}\,dh}{\lambda\sinh(a/2\lambda)(1-4h/a)^{1/2}}. \qquad \ldots\ldots(23)$$

Remembering that h cannot exceed the height $\tfrac{1}{4}a$ of an arch, the fraction of free distances in the direction of the electric field exceeding h is

$$\int_h^{a/4} P(h)dh = \frac{\sinh\{(a/2\lambda)(1-4h/a)^{1/2}\}}{\sinh(a/2\lambda)} \qquad \ldots\ldots(24)$$

It is not difficult to show that this expression reduces to the form $\exp(-l/\lambda)$ as the magnetic field $H\to 0$, i.e. as $a\to\infty$ and $h\to l$. Consequently, if we put λ'' for λ and W''/E for h, where λ'' and W'' are the values appropriate to this case, we obtain an expression for the fraction of electrons colliding with energy in excess of eW'', i.e. the fraction of ionizing collisions, of the form

$$\sinh\{(a/2\lambda'')(1-4W''/Ea)^{1/2}\}/\sinh(a/2\lambda'') \qquad \ldots\ldots(25)$$

which reduces to $\exp(-W''/E\lambda'')$ (cf. eqn. (18)) when $H=0$. Hence we must have, from eqn. (20),

$$W''/\lambda'' = W'/\lambda' = Bp\,. \qquad \ldots\ldots(26)$$

To obtain α we must multiply expression (25) by the number of collisions per unit length in the direction of the electric field, a quantity which we may denote by $1/\bar{h}''$. Doing this, and using (26), we obtain

$$\alpha = (1/\bar{h}'')\sinh\{(a/2\lambda'')(1-4Bp\lambda''/Ea)^{1/2}\}/\sinh(a/2\lambda'') \qquad \ldots\ldots(27)$$

as the equation equivalent to (18).

We now encounter a difficulty, in that we lack criteria which will allow us to determine λ'' and $\bar{h}''$. Just as, in eqn. (18), the identification of λ' with the kinetic theory mean free path, or the mean free path measured in other ways, would lead to values of λ not in accord with observation, so, in eqn. (27), the identification of λ'' and $\bar{h}''$ with similar quantities measured or calculated in other ways may prove unsatisfactory.

It might seem logical to relate $\bar{h}''$ directly to λ'' by an equation of the form

$$\bar{h}''/a = F(\lambda''/a), \qquad \ldots\ldots(28)$$

similar to eqn. (16*b*). This leads to the equation

$$\alpha = \{1/aF(\lambda''/a)\}\sinh\{(a/2\lambda'')(1-4Bp\lambda''/Ea)^{1/2}\}/\sinh(a/2\lambda'') \qquad \ldots\ldots(29)$$

which reduces, when $H \to 0$, to $\alpha = (1/\lambda'')\exp(-Bp/E)$, so that, by comparison with (18) and (20), we must put $\lambda'' = \lambda'$.

Some calculations have been made of sparking potentials based on use of eqn. (29), with $\lambda'' = \lambda'$, and including the correction for electron recapture by the cathode. Compared with Penning's observed values, the calculated minimum sparking potentials, and the magnetic fields at which they occur, are too small by an order of magnitude, and calculations based on eqn. (18) are, in fact, in much better agreement with observation than calculations based on eqn. (29). It does not seem, therefore, that eqn. (29) is a satisfactory expression for α.

We may approach the question of the value to be given to $\bar{h}''$ in another way. To bring eqn. (18) (for the non-magnetic case) into line with observation it is necessary to let the mean free distance λ' travelled in the direction of the electric field differ by a certain factor from λ, the mean free path defined in other ways; $\lambda'/\lambda = L'/L = 1/AL$, of the order of two or three. When a magnetic field is present, we may make a similar correction, by taking $\bar{h}''$ to be equal to $1/AL$ times $\bar{h}$, the mean free distance in the direction of the electric field, calculated from λ. Thus if, in (27), we put $\bar{h}'' = \bar{h}(\lambda)/AL = aF(\lambda/a)/AL$ and $\lambda'' = \lambda$, we get

$$\alpha = \{AL/aF(\lambda/a)\}\sinh\{(a/2\lambda)(1-4BL/Ea)^{1/2}\}/\sinh(a/2\lambda) \qquad \ldots\ldots(30)$$

A comparison of the values of α given by eqns. (18) and (30) shows that, if $4BL/Ea < 0{\cdot}9$, they never differ by more than about 50%, so that under these circumstances, using (30) instead of (18) will not make any major difference to the sparking potential curves. In fig. 4 the line corresponding to $BL/Ea = 1$ is shown dotted. (From (3), $4BL/Ea = BLeH^2/2E^2\mathbf{m} \times 10^8 = 0{\cdot}84(H/E)^2$, so that this line also corresponds to $H/E = 1{\cdot}1$.) Since, according to eqn. (30), α is zero at all points on and below this line, the sparking potentials calculated using eqn. (30) must lie above it.

It seems then that the use of eqn. (30) will result in sparking potential curves in general qualitative agreement with observation, but not greatly different from the curves of fig. 4. It must be remembered that many arbitrary assumptions underlie eqn. (30), and it is impossible to say to what extent it is a valid approximation to α until measurements of α in a magnetic field are available with which it may be compared.

Whatever its range of validity may be, it will certainly not be valid when $H > E$. In this region our basic assumptions are no longer reasonable approximations to the truth, for clearly, when electrons cannot attain from rest in the electric field sufficient energy to ionize, our assumption of completely inelastic collisions is seriously astray. A theory satisfactory under these conditions must take account of elastic collisions, and the possibility of an electron gaining sufficient energy to ionize as the result of two or more favourable collisions. This would involve an analysis of the situation in which the energy of random motion of the electrons is comparable with the energy acquired from the field between collisions.

This region, in which H/E is relatively high, is just the region where the calculated potential curves fail to show even qualitative agreement with observation, for Penning finds that at high magnetic fields the sparking potential decreases again, while all our calculations give a steady rise.

It seems that a better understanding of the nature of the observed sparking potential curves must wait on experimental measurement of α as a function of E, H and p. If such measurements were available it might be possible to find an empirical analytical expression for α to replace Townsend's relation or, failing that, to work graphically from the experimental curves.

ACKNOWLEDGMENTS

I wish to acknowledge the benefit received from discussions with Professor V. A. Bailey and the assistance of Mr. J. T. Bishop, who helped with, and devised means of shortening, the computational work.

APPENDIX

To show that $x^{-1}\coth x - x^{-2}$ is a monotonic function of x for $x>0$.

Let $y = x^{-1}\coth x - x^{-2}$,

then

$$\begin{aligned} dy/dx &= -x^{-2}\coth x - x^{-1}\operatorname{cosech}^2 x + x^{-3} \\ &= (2\sinh^2 x - x\cosh x \sinh x - x^2)/(x^3\sinh^2 x); \end{aligned}$$

hence, when $x>0$, dy/dx has the same sign as

$$\begin{aligned} \cosh 2x - 1 - \tfrac{1}{2}x\sinh 2x - x^2 &= \sum_{n=1}^{\infty}\frac{(2x)^{2n}}{(2n)!} - \tfrac{1}{2}x\sum_{n=1}^{\infty}\frac{(2x)^{2n-1}}{(2n-1)!} - x^2 \\ &= \sum_{n=2}^{\infty}\frac{(2x)^{2n}}{(2n)!} - \tfrac{1}{4}\sum_{n=2}^{\infty}\frac{(2x)^{2n}}{(2n-1)!} = \sum_{n=2}^{\infty}(2x)^{2n}\left\{\frac{1}{(2n)!} - \frac{\tfrac{1}{4}}{(2n-1)!}\right\}. \end{aligned}$$

But $4(2n-1)! \leqslant (2n)!$ for $n \geqslant 2$, so that $dy/dx<0$ when $x>0$ and y is monotonic if $x>0$.

REFERENCES

von Engel, A. and Steenbeck, M., 1932, *Electrische Gasentladungen*, Band I (Berlin : Springer), pp. 39, 118.

Kugler, G. A., Ollendorf, F., and Roggendorf, A., 1933, *Z. Phys.*, **81**, 733.

Meyer, E., 1919, *Ann. Phys., Lpz.*, **58**, 297.

Penning, F. M., 1937, *Physica*, **4**, 71.

Townsend, J. S., 1915, *Electricity in Gases* (Oxford : University Press), p. 295.

Valle, G., 1950, *Nuovo Cim.*, **7**, 174.

Wehrli, M., 1922, *Ann. Phys., Lpz.*, **69**, 285.

Metal Transfer during Static Loading and Impacting

By E. RABINOWICZ*
Research Laboratory on the Physics and Chemistry of Surfaces,
Department of Physical Chemistry, Cambridge

MS. received 19*th February* 1952

ABSTRACT. Experiments are described which investigate the metal transfer which takes place when a radioactive hemispherically ended slider is pressed normally into a flat surface of the same or another metal.

It is found that metal transfer in the form of a number of small fragments is always observed, and this supports the view that metallic junctions are formed between metal surfaces in contact. Experiments in which the load is varied suggest that at the higher loads the oxide layer on the metal surfaces is broken up to a larger extent than at lower loads, and a more than proportional increase in metallic interaction and transfer takes place. An analogous effect is observed in the presence of boundary lubricants.

When the surfaces are impacted together, very similar results are obtained. Somewhat less pick-up is observed than for static loading, and the difference is probably due to the fact that it takes time for strong junctions to be formed. Impact experiments with surfaces covered by lubricants show that a lubricant layer may be trapped between the surfaces, and this produces a large reduction in pick-up without greatly reducing the amount of plastic deformation.

§1. INTRODUCTION

The theory of Bowden, Moore and Tabor (1943) on the nature of the friction force postulates that, when clean metal surfaces are pressed together, metallic junctions are formed between them and these have to be sheared before sliding can occur.

A direct test of this theory would consist of a measurement of the adhesion of clean metal surfaces pressed normally together. However, in general no such adhesion is observed, because when the force pressing the surfaces together is removed, a release of elastic stresses occurs and the junctions are broken one by one. In the case of soft metals such as indium these elastic stresses are small, and McFarlane and Tabor (1950) and Parker and Hatch (1950) have recently shown the considerable adhesion of clean indium surfaces.

Another method of investigating these metallic junctions, applicable to hard as well as to soft metals, consists in making one of the metal surfaces radioactive and then, after pressing the surfaces together, measuring the radioactivity of the other. The transfer of radioactivity under these conditions gives direct proof of the formation of metallic junctions.

A point of some interest is the time factor when the surfaces are pressed together. Crook and Hirst (1950), working with indium, have shown that the force of adhesion is much smaller for impacted surfaces than for surfaces pressed slowly together.

Experiments were therefore carried out in which a metal slider with a hemispherical end, previously made radioactive, was pressed into a flat surface of the same or another metal, slowly or under impacting conditions, and the transfer of activity to the other surface investigated.

---

* Now at the Lubrication Laboratory, Department of Mechanical Engineering, Massachusetts Institute of Technology.

§ 2. EXPERIMENTAL

In preliminary experiments the hemispherical slider was fixed into a simple holder and pressed into the other surface by a spring. The apparatus was not very rigid and no precautions were taken to eliminate tangential motion as the two surfaces came into contact. Under these conditions it was found that some lateral motion—estimated at 10^{-4} to 10^{-3} cm—occurred, and this had a profound effect on the amount of radioactive transfer, so that later experiments were carried out with a much more rigid apparatus. A special holder was made which fitted a Vickers Hardness Tester, and the hemispherically ended slider then occupied the position normally taken up by the diamond indenter. Loads of 1–100 kg could be applied, and the time of loading was normally 10 seconds.

For the impact experiments the slider was screwed into the bob of a compound pendulum (fig. 1). Impact energies of 10^4 to 10^6 ergs could be used. The time of impact was estimated (cf. Bowden and Tabor 1941) as about 10^{-4} sec.

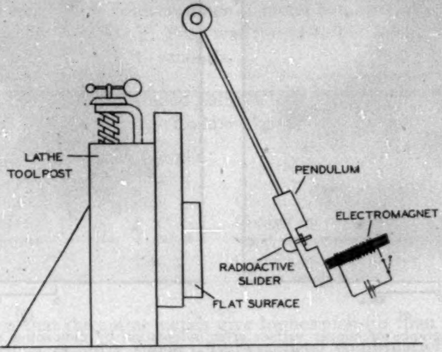


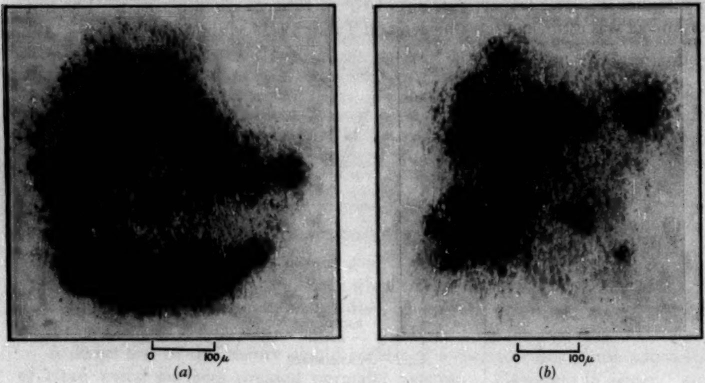
Fig. 1. Sketch of apparatus used for the impact experiments.

Before the experiments the surfaces were cleaned to remove surface contamination, the radioactive hemispherical slider by rotation in the shaft of an electric motor whilst pressing fine carborundum paper against its surface and the flat surface by lapping on fine carborundum paper under water, followed by drying on clean filter paper.

The radioactive sliders were prepared by irradiating sliders of the required shape in the B.E.P.O. at Harwell. The amount and nature of the resultant activity varies widely for different metals. Sliders made of copper, cadmium, platinum and stainless steel (18% chromium) have proved suitable.

The radioactivity transferred on to the flat surface after indentation experiments had been carried out was measured by placing the flat surface on a sheet of photographic film, Ilfex x-ray film being normally used. After development of the film, a blackened region corresponding to the transferred radioactivity could be seen. The method was made quantitative by visual comparison of the amount of blackening with that produced when a thin foil of the same metal, irradiated together with the slider and thus of the same specific activity, was placed on a piece of the same film for known periods of time. The amount of blackening was found to depend only on the total amount of radiation received by the film, not on the time or intensity separately, i.e. the reciprocity law was obeyed (Rabinowicz 1950).

The sensitivity of the method is such that, in the case of copper, metal transfer of 10^{-12} g per impression may be fairly readily detected. The accuracy of the method depends on whether it is required to determine the amount of pick-up absolutely or to compare two impressions made at the same time. In the former case the probable error is about 30%, in the latter case 15%. The reproducibility of normal loading experiments is fairly good, results obtained in different experiments, nominally the same, showing mean variations of about 40%.



(*a*) (*b*)

Fig. 2. Autoradiographs of copper surface after indentation by radioactive copper hemisphere under a load of 10 kg.

(*a*) Clean. Metal transfer $= 6 \times 10^{-16}$ g.
(*b*) Surface covered with cetane. Metal transfer $= 10^{-10}$ g.

(Exposures have been adjusted to give equal intensities of blackening.)

§ 3. RESULTS OF STATIC LOADING EXPERIMENTS

(i) *Effect of Lateral Movement*

Experiments made with the first, not very rigid, apparatus showed the great importance of eliminating lateral movement when bringing the two surfaces together and removing them. In a typical case, copper pressed into copper, clean, under a load of 4 kg, the amount of transfer varied between 3×10^{-8} and 3×10^{-9} g, the exact value depending probably on the amount of lateral movement whilst making the impression. When great care was taken to eliminate this motion consistent values of 10^{-10} g were obtained.

This value was also obtained when experiments were carried out under the same conditions in the Vickers Hardness Tester, in which lateral motion is extremely small, and thus may be considered with some confidence as that for pure normal loading.

(ii) *Nature of the Metal Transfer*

Figure 2 (*a*) shows an autoradiograph produced by clean copper pressed into copper using a load of 10 kg. It is seen that metal transfer has taken place as a number of fragments, each of which has produced a blackened zone.

The total amount of pick-up for this impression was 6×10^{-10} g. Assuming that most of this pick-up is contained in about 20 fragments equal in size, we can deduce for each (if spherical) a diameter of 2×10^{-4} cm, or 25 000 Å. This shows that it is probably metallic copper and not just the oxide layer that has been transferred, since the latter for copper is at most a few hundred Ångströms thick (Evans 1946), and a flat flake of oxide of the required mass would have an area bigger than the size of the metallic junctions. At 4 kg the fragment diameter is 10 000 Å, and the same considerations apply.

(iii) *Use of Different Metals and its Effects*

Impressions were made for a number of metal combinations, and the results are collected in table 1. The pick-up was found to be independent of the time of loading within the range 1–50 seconds.

Table 1. Metal Transfer produced using Clean Surfaces of the Metals stated.
Load = 4 kg

Metals	Metal transfer (10^{-10}g)	Metals	Metal transfer (10^{-10}g)
Cadmium on cadmium	2	Cadmium on steel	3
Copper on copper	1	Copper on steel	0·2
Platinum on platinum	4	Steel on copper	1
Steel on steel	1	Copper on cadmium	10
		Platinum on silver	15

It will be seen that the softer metals give higher pick-up than the harder ones, although this effect is fairly small. Unlike metal combinations give as much pick-up as do like ones.

(iv) *Repeated Impressions and their Effect on the Pick-up*

The final finish given to the top surface was produced by rotating it against 600-grade carborundum paper, and thus a surface with small but sharp asperities was produced. It was to be expected that the amount of pick-up would be large until these asperities had been worn off, and this behaviour was found. The first impressions produced more metallic transfer than subsequent ones; after the fifth impression the amount of pick-up was constant, as is shown in table 2. Each impression was made on a different part of the lower surface.

Table 2. Clean Copper on Copper. Load = 20 kg

Impression	1st	2nd	3rd	4th	5th	6th	7th	8th
Transfer (10^{-10}g)	600	250	150	100	100	50	50	50

In carrying out experiments it is important that sufficient impressions are taken to eliminate this initial effect. All other values in this paper correspond to the steady-state value (50×10^{-10} g in table 2).

(v) *Effect of Surface Roughness*

Experiments were carried out using finishes of different grades of carborundum paper for the lower surface. The results are shown in table 3 (the roughness was measured by profilometer).

Table 3. Clean Steel on Copper. Load = 20 kg

Finish (grade)	600	220	80
Roughness (microinch)	1	2	10
Pick-up (10^{-10}g)	50	60	70

It will be seen that the changes in roughness have had very little effect on the amount of pick-up in the steady state; the differences actually found are so small as to be only barely significant.

(vi) *Variations in the Load and its Effects*

Experiments were carried out in which the load was varied from 1 to 100 kg. In all cases an arrangement of a hard indenter and a softer lower surface was used, since otherwise the geometry of the indenter would have changed during the course of the experiment.

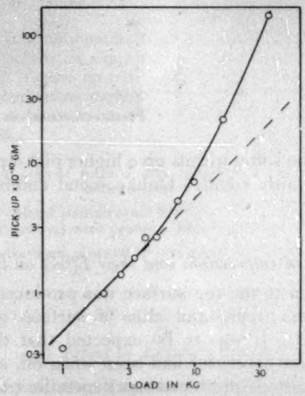


Fig. 3. Metal transfer as a function of load for a clean copper indenter and an annealed copper surface.

Figure 3, for hard copper on annealed copper, clean, shows that at low loads the pick-up is proportional to the load, but at high loads deviations occur. Similar data were obtained with the other metal combinations, and the effect appears to be a general one. All the experimental results obtained could be fitted to curves of the type Pick-up $= c_1L + c_2L^2$, where c_1 and c_2 are constants and L is the load.

(vii) *Use of Lubricants*

It was found that use of a lubricant reduced the amount of metal transfer, as may be seen in table 4.

Table 4. Copper on Copper : Amount of metal transfer in 10^{-10} g

Load (kg)	4	10	50
No lubricant	1	6	150
Cetane	0·06	1	50
Solution of stearic acid in cetane	0·01	0·25	15

The solution of stearic acid, known to be a better boundary lubricant than cetane itself in the case of sliding surfaces, produces less pick-up also in the case of stationary surfaces. Both lubricants lose much of their effectiveness at the higher loads.

Figure 2 (*b*) shows that the use of a lubricant has not greatly reduced the number of metal fragments transferred. It is the average size of the fragments that has been decreased.

§ 4. EXPERIMENTS WITH IMPACTED SURFACES

Experiments were carried out to check whether the behaviour of the impacted surfaces was the same as that of the statically loaded ones as described in the preceding sections. Very similar results were obtained. For example: for a number of impressions at a constant impact energy the metal transfer decreased for the first four or five impressions and then became constant, in agreement with the results shown in table 2; when the impact energy was varied it was found that at low energies the pick-up was proportional to the energy of impact, but at high energies deviations occurred and the pick-up was greater than would be given by a proportionality relationship (cf. fig. 3).

The important experiment, however, was a quantitative comparison of the amount of metal transfer with that obtained under static loading.

(i) *Comparison of the Metallic Transfer for Statically and Kinetically Loaded Surfaces*

To test how the amounts of pick-up compared in the two cases, loads and impact energies were selected such that indentations of equal size were produced. Table 5 shows the results obtained.

Table 5. Hardened Copper on Annealed Copper. Clean

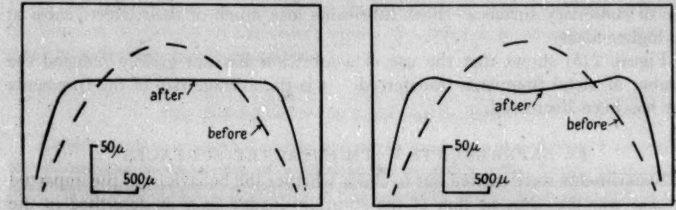
Diameter of impression (mm)		0·4	0·8	1·2
Mass of pick-up (10^{-10} g)	impacted surfaces	3	20	100
	static loading	5	30	100

The amount of metal transferred is approximately the same in the two cases. Static loading produces somewhat more pick-up, but the difference is small.

(ii) *Impact Experiments with Lubricated Surfaces*

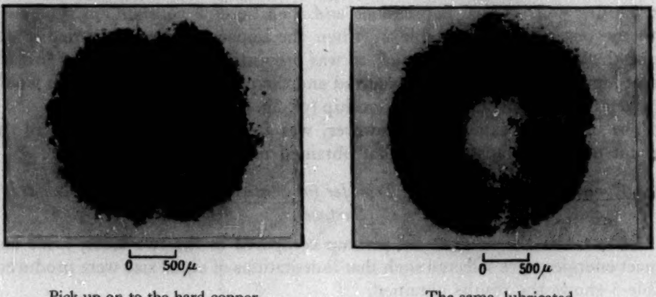
Experiments were carried out with surfaces covered by various paraffin oils and a range of silicone oils. It was found that metal transfer still took place, but on a reduced scale compared with that for clean surfaces. The reduction was in two parts; firstly, there was an overall decrease in pick-up, similar to that observed during static loading in the presence of a lubricant; secondly, there was another effect as a result of which no metal transfer took place at the centre of the indentation. Figure 4 shows the results of an experiment in which radioactive annealed copper hemispheres impacted hard copper surfaces, both clean and covered with a viscous silicone oil. The profile of the indenter as measured by a Talysurf profilometer and an autoradiograph of the flat surface after impact are shown in each case. As a result of the use of the lubricant, metallic contact at the centre of the indentation has been prevented and additional deformation produced.

Further experiments were carried out in which hardened steel balls of 2·5 cm diameter were impacted into copper surfaces. In this case the metallic interaction was studied by measurement of the electric conductance between the steel and the copper during impact. With clean surfaces the electrical measurements showed that, as expected, metallic contact occurred at impact. When an oil of viscosity 10 centistokes was spread on the surface, the collision process showed the same characteristics. Deformation of the surface took place and metallic contact occured. When the surface was covered with an oil of viscosity 500 centistokes a marked change was observed. No metallic contact or break-through occurred during the impact, but the copper surface still showed very marked deformation. Typical indentations are shown in fig. 5. It will be seen that the deformation produced in the presence of the oil is somewhat different in shape from that obtained with clean surfaces. It is conical rather than spherical, with a marked 'pip' or dimple at the centre.



Slider of annealed copper after impacting a hard copper surface. Clean.

The same, but surface covered with oil of viscosity 1000 cst.



Pick-up on to the hard copper.

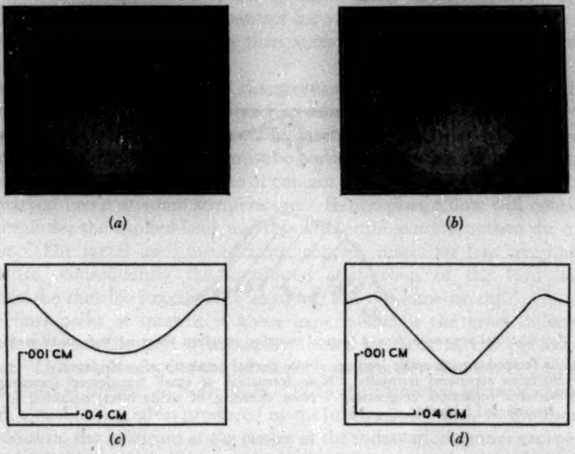
The same, lubricated.

Fig. 4.

§ 5. DISCUSSION

These experiments on the pressing of a hemispherical slider into a flat metal surface show that, under very varied conditions, metal fragments are transferred from one surface to the other, and this shows in a direct way that strong junctions are formed between metal surfaces in contact.

The amount of metal transferred is small. In a typical case, copper on steel, clean, load 4 kg, the pick-up is 2×10^{-11} g, which corresponds to a uniform layer of thickness 0·1 Å over the area of the indentation. This value may be compared with that for sliding surfaces. For copper on steel, clean, load 4 kg, Rabinowicz and Tabor (1951) have found metal transfer equivalent to a uniform layer of thickness 500 Å over the sliding track, i.e. there is a difference by a factor of 5 000 between the two cases. It follows that even small amounts of lateral motion while the static indentations are being produced may lead to comparatively large amounts of metallic transfer, far greater than for purely static loading. For example, a lateral movement of only 10^{-3} cm may lead to a hundred-fold increase of pick-up over that for pure static loading.



(*a*) (*b*)

(*c*) (*d*)

Fig. 5. Micrographs and profiles of copper surface after impacting by a steel hemisphere. (*a*) and (*c*), clean; (*b*) and (*d*), surface covered with a silicone oil of viscosity 500 cst.

We must now consider the fundamental differences between the processes of sliding and of static loading. When the metal surfaces are brought into contact, junctions (usually referred to as 'metallic junctions') are formed between the prominent asperities (Bowden, Moore and Tabor 1943). Each junction will consist partly of a region where oxide layers still cover each metal and constitute the interface, partly of a region where the oxide layer has been broken and contact is established between the bare metals. Only in this latter region is there any likelihood that the strength of the junction is as great as that of the bulk metals and, since this is a prerequisite for the transfer of metal fragments, only in this region can we expect the occurrence of pick-up. It follows that the fragments formed during static loading will have dimensions smaller than that of the junction (fig. 6).

If, however, the metallic junction is broken by shear, as during sliding, conditions are different. In this case, owing to the different geometry, the breaking of the junction is preceded by a shearing action which will tend to break up the oxide layer and produce metallic contact over the whole of the junction. Then, if a fragment is produced, its dimensions will be those of the junction. In addition, the whole junction may increase in size before shearing takes place (McFarlane and Tabor 1950).

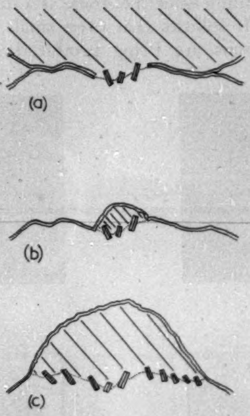


Fig. 6. Hypothetical appearance of a typical metallic junction showing the oxide surface layer.

(*a*) As formed during static loading. Note partial break-up of oxide layer.
(*b*) Surfaces separated normally. Note formation of small transferred fragment.
(*c*) Surfaces separated tangentially. Note shearing of oxide layer, resulting in a large transferred fragment.

Considering next the variation of pick-up with load, it is found that the pick-up increases more than proportionately at the higher loads. This is probably due to the fact that the more severe deformation of the surfaces at the higher loads leads to a more widespread breakdown of the oxide layer, and thus metal-to-metal contact is made over a larger fraction of the true area of contact, which in turn produces a greater amount of metal transfer. An analogous effect has been observed in the case of sliding surfaces by Whitehead (1950), who observed the increase of coefficient of friction with load within a certain range of loads and was able to show that this did correspond to different degrees of breakdown of the oxide layer.

The behaviour in the presence of a lubricant is consistent with our present views on the nature of lubrication. Boundary lubricants act by reducing the area over which metallic contact occurs. Thus in the presence of the lubricant some metallic transfer occurs, though less than in the absence of the lubricant.

The presence of the lubricant diminishes the size, rather than the number, of metallic junctions. In these experiments we find that, at higher loads, the lubricants are less effective in reducing the amount of pick-up, and this may be due to the fact that, as in the presence of the oxide layer considered previously, the lubricant layer is broken up at excessive deformation of the surface (McFarlane and Tabor 1950), and as a result more metal-to-metal contact occurs. Again the results are analogous to those for sliding surfaces, where it is found that in the presence of good boundary lubricants the amount of metallic transfer increases more than proportionately with the load (Rabinowicz and Tabor 1951).

From the fact that impacted surfaces—in which the time of contact is of the order of 10^{-4} sec—give nearly the same value of pick-up as do surfaces in contact for times in the range 1 to 50 seconds, it appears that the formation of metallic junctions takes place very rapidly. However, the results show that the metallic transfer is somewhat less with impacted surfaces—by about 25%. This fits in well with the results of Sampson, Morgan, Reed and Muskat (1943), who found that surfaces in stationary contact for very short times gave lower values of the static coefficient of friction than surfaces that had been in contact for some seconds.

At first sight the results are in disagreement with those of Crook and Hirst (1950) who, working with indium and measuring the force of adhesion, found that impacted surfaces gave less metallic interaction than statically loaded ones by an order of magnitude. But it must be borne in mind that the greater strength of junctions formed over a long time of contact is probably due to creep. Indium shows marked creep at room temperature. Hence plastic flow will continue to take place under the applied load, and this will continue to strengthen the metallic junctions. The metal used in our tests, copper, shows far less creep at room temperature, consequently the continued application of the load will not strengthen the metallic junctions to anything like the same extent.

The final point of interest in these experiments is the great difference in behaviour of statically and dynamically loaded surfaces in the presence of a lubricant. One result of using the lubricant is an overall reduction in the pick-up, and this effect is present in both cases. With the impacted surfaces there is, in addition, a cushioning effect produced by the lubricant. Owing to the short time of the collision, the lubricant at the centre of the indentation cannot escape and is trapped, and this in turn leads to the absence of pick-up at the centre of the indentation. Under suitable conditions metallic contact may be completely eliminated. However, deformation of the metals still takes place, and the indentation is characterized by a 'pip' or dimple at the centre corresponding to the trapped lubricant.

These results have important technological significance. They show that in cases where the loading is impactive (e.g. in gear teeth), the presence of the lubricant, while reducing, or even eliminating, metallic contact, may have little effect on the amount of deformation produced in the metal surfaces.

ACKNOWLEDGMENTS

It is with pleasure that I thank Dr. F. P. Bowden for his interest in this work, Dr. D. Tabor for his constant encouragement and advice, and the Department of Scientific and Industrial Research for a maintenance grant.

REFERENCES

Bowden, F. P., Moore, A. J. W., and Tabor, D., 1943, *J. Appl. Phys.*, **14**, 80.
Bowden, F. P., and Tabor, D., 1941, *J. C.S.I.R., Austr.*, **14**, 152.
Crook, A. W., and Hirst, W., 1950, *Research*, **3**, 432.
Evans, U. R., 1946, *Metallic Corrosion, Passivity and Protection* (London : Arnold).
McFarlane, J. S., and Tabor, D., 1950, *Proc. Roy. Soc.* A, **202**, 224.
Parker, R. C., and Hatch, D., 1950, *Proc. Phys. Soc.* B, **63**, 185.
Rabinowicz, E., 1950, *Proc. Phys. Soc.* A, **63**, 1040.
Rabinowicz, E., and Tabor, D., 1951, *Proc. Roy. Soc.* A, **208**, 455.
Sampson, J. B., Morgan, F., Reed, D. W., and Muskat, M., 1943, *J. Appl. Phys.*, **14**, 689.
Whitehead, J. R., 1950, *Proc. Roy. Soc.* A, **201**, 109.

Further Experiments on Evaporation from Small, Saturated, Plane Areas into a Turbulent Boundary Layer

By D. R. DAVIES and T. S. WALTERS
Mathematics Department, University of Sheffield

MS. received 3rd March 1952

ABSTRACT. This paper describes an experimental investigation designed to evaluate the effect of width and length on rates of evaporation from small, saturated, plane, rectangular areas placed in a turbulent boundary layer. An approach is made to the problem of estimating the distribution of temperature on the evaporating surface by using a number of thermojunctions situated at various points under the surface. The higher rate of evaporation near the edges of the area is well illustrated by this method, greater temperature differences between the surface and air stream being observed near the edges due to greater latent heat effects. Provided the crosswind width of the evaporating area is not less than about half the downwind length, Sutton's two-dimensional solution of this limited area problem is found to lead to tolerable accuracy in the circumstances tested.

§ 1. INTRODUCTION

In a previous paper (Davies and Walters 1951) the authors have described experiments, carried out in an open circuit wind tunnel, designed to test the dependence of rate of evaporation on width of area. These experiments have now been extended to areas of length up to 15 inches and widths varying from 1 to 6 inches. The technique followed was again the 'gravimetric' method, which has been previously described fully (Pasquill 1943). In order to make as close a comparison as possible between theory and experiment for absolute values of rates of evaporation a good estimate is needed of the saturation vapour concentration at the surface. This is accomplished by soldering thermojunctions at various points on the lower side of a thin brass plate, on the upper face of which the saturated pieces of filter paper are placed. The temperature distribution and hence the surface vapour concentration distribution is then determined by connecting the thermojunctions to a calibrated potentiometer. The temperature differences between the air stream in the tunnel and the evaporating surface, due to the latent heat of vaporization of the liquid, are found to be appreciably higher near the edges of the area than elsewhere inside the area. This gives further confirmation of the authors' previous suggestion (Davies and Walters 1951) that rate of evaporation is greater near the edges of the area.

The greater range of length and width now tested enables an interesting comparison to be made with experiment of Sutton's theoretical results on evaporation (Sutton 1949) and the solution obtained by the present authors (Davies and Walters 1951). The former treatment neglects the enhanced rate of evaporation near the longitudinal edges and the latter assumes the rate of evaporation to be independent of the downwind distance from the leading edge.

§ 2. EXPERIMENTAL ARRANGEMENTS

The general method adopted for measuring rates of evaporation followed that already described in previous work (Davies and Walters 1951), except that the new glass plate used was about 17 inches long and 8 inches wide, and a closer approach to the temperature of the evaporating surface was attempted. The width of the plate was limited by the dimensions of the arm of the available balance. The length of the plate was determined by the length of the flat portion of the faired holder, into which a recess of the necessary dimensions was made.

When the required evaporation observations had been completed the glass plate was replaced by a thin brass plate of the same area and 0·04 in. thickness. Thermojunctions were soldered to the underside of the plate at the points marked A, B, C, D, E, F in fig. 1, and temperatures measured by means of a calibrated potentiometer.

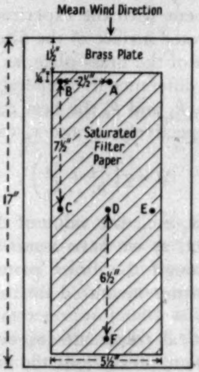


Fig. 1. Diagram indicating positions of thermojunctions at A, B, C, D, E, F.

A rectangular piece of filter paper, 13½ in. long and 5½ in. wide, was placed on the brass plate in the holder in the position shown in fig. 1, and saturated with aniline, with the holder resting centrally in position on the floor of the working section of the wind tunnel. The air speed in the tunnel was adjusted and maintained at the constant value, 1 330 feet per minute, used during the evaporation tests. Readings were made of the thermojunction temperatures (at points A, B, C, D, E, F), the air temperature in the tunnel stream, and the air temperature in the neighbourhood of the potentiometer, which was sheltered as much as possible and set just outside the tunnel door. After five or six minutes the readings settled down to steady values until the filter paper began to dry, which took about forty minutes. These temperature results together with the new evaporation results are given in the following section.

§ 3. EXPERIMENTAL RESULTS AND DISCUSSION

The temperature measurements are discussed first, as they lead to a correction factor which must be evaluated in order to compare theory and experiment for absolute magnitudes of rates of evaporation. During the experiment the temperature of the air stream in the tunnel remained constant and was found to be 0·5° c cooler than room temperature near the potentiometer. After this difference was allowed for, the temperature differences between the tunnel air stream and the evaporating surface were evaluated. These were 1·7° c for positions A, B, C, E, 1·4° c for D and 1·8° c for F.

The temperature decreases near the edges (A, B, C, E, F) of the area are observed to be about 20% greater than in the centre (D) of the area: this indicates a considerably greater rate of evaporation near the edges than at points in the interior of the area, and confirms the result of a previous paper (Davies and Walters 1951). It is interesting to note that, in the *special circumstances* of this previous work, the theoretical variation of rate of evaporation across an evaporating area is proportional to $y^{-m/2}(w-y)^{-m/2}$, where w is the width of the area, y the distance of the point considered from an edge, and m is a turbulence parameter, whose value was found (Davies and Walters 1951) to be 0·15. This formula predicts that, at a point situated ¾ in. (approximately the position of thermojunctions C and E) from an edge of an area 5½ in. wide, the rate of evaporation is about 15% higher than in the centre. This is in reasonable agreement with the experimental value of about 20%.

The temperature of the outer surface of the filter paper is assumed to be the same as that of the underside of the brass plate, as given by the thermojunction readings. If Δt is the actual difference in temperature between the two surfaces concerned, q the flux of heat, k_1 and k_2 the heat conductivities of the brass and filter paper–aniline layers respectively, x_1 and x_2 the thicknesses of these layers, then

$$\Delta t = q\left(\frac{x_1}{k_1}+\frac{x_2}{k_2}\right). \qquad \ldots\ldots(3.1)$$

From measurements of x_1, x_2 etc., the value of Δt is found to be 3×10^{-6}°c, and this is a maximum effect, as we have assumed that all the latent heat of evaporation is supplied through the brass plate. This result justifies the assumption that the temperatures measured are those of the actual evaporating surface.

It is assumed in theoretical treatments of evaporation problems (Sutton 1949) that the saturated vapour pressure of the evaporating liquid is constant over the area and consequently the temperature of the surface should strictly be kept constant over the area. However, although the temperature *decreases* in the present work varied appreciably over the surface, the *actual* temperatures could be regarded as effectively constant. In the conditions considered the surface temperature in the centre was 17·1° c and along the edges it was 16·8° c, the tunnel air temperature being 18·5° c. At this temperature a suitable mean reduction factor thus appears to be 1·5° c. If 17·0° c is taken as a mean surface temperature, the variation in actual temperature on the plate is extremely small, and the consequent variation of saturated vapour pressure is also small (from the appropriate curve for aniline), and so the assumption of constant surface temperature leads to little error. It is reasonable to suppose that the factor of 1·5° c will not vary significantly over the small range of room temperatures considered in the evaporation tests.

Several measurements of rates of evaporation have been made which are additional to the results already given in a previous paper (Davies and Walters 1951). For a constant length l of area and a constant mean air speed, log E was plotted against log w, E being the rate of evaporation in lb. per minute and w the width in inches. The results are shown in fig. 2 in the form of two sets of points; the conditions tested are described in the figure. Twenty-two points are plotted in each set and lines of regression are computed and drawn. The slopes are found to be 0·84 and 0·86, in good agreement with the theoretical result of 0·85, described previously (Davies and Walters 1951).

Measurements to determine the dependence of E on l were also made in the conditions detailed in fig. 3: the points corresponding to the lengths 3·5, 5·0, 13·0, 15·0 in. (marked as full circles in fig. 3) were obtained from lines of regression giving the dependence of E on w (see fig. 2). From fig. 3, E is found to be proportional to $l^{0\cdot76}$. The result given by Sutton's well-known theory (Sutton 1949) which assumes unlimited width of area, is E proportional to $l^{(1+m)/(1+2m)}$, where m is a parameter determined from the velocity profile in the turbulent boundary layer over the area. As described previously (Davies and Walters 1951) the evaporating area tested was positioned in a *fully turbulent boundary layer*, and a value for m of 0·15 was obtained. This leads to a theoretical value for $(1+m)/(1+2m)$ of 0·88. The consequences of this difference in theoretical and measured dependence on length is clearly seen in the table.

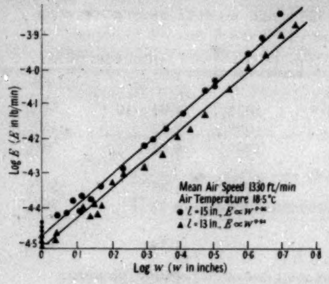


Fig. 2.

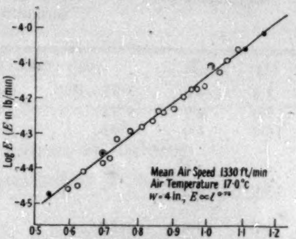


Fig. 3.

The considerable range of length and width of area now tested experimentally makes a comparison of available theoretical treatments interesting. In the table some of the results of the evaporation measurements are given together with corresponding theoretical values computed, using Sutton's (1949) formula

$$E = KU_1^{(2-n)/(2+n)} l^{(1+m)/(1+2m)} w, \qquad \ldots\ldots(3.2)$$

and the formula obtained by the present authors (Davies and Walters 1951)

$$E = \frac{2\pi\chi_0 a_z U_1^{1-n} f^{m/2}\Gamma(1-m/2) l w^{1-m}}{\Gamma(m/2)\Gamma(\frac{3}{2}-m/2)\Gamma(\frac{1}{2}-m/2)}, \qquad \ldots\ldots(3.3)$$

where K has been tabulated (Pasquill 1943), χ_0 is the saturation vapour concentration at the temperature of the surface, a_z and n have been specified in previous work and are easily calculated from the observed value of m, and U_1 is the mean air speed at a convenient reference height. The f factor in eqn. (3.3) has been discussed previously (Davies and Walters 1951) and enters only to small power, $m/2$. Near the surface in the earth's atmosphere it has been suggested (Davies and Walters 1951) that an appropriate value, in adiabatic conditions, is 1·6. Some recent results (Townsend 1951) on the variation of eddy velocities with distance from the wall in a turbulent boundary layer appear to some extent to be relevant here. Assuming that the results obtained by Townsend apply in the wind tunnel used by the authors, a value for f of 1·0 is computed at a height of about 0·4 in. over the evaporating surface and of 1·6 at a height of about 0·2 in., the value increasing as the surface is approached. A mean value for f is needed over the height of the *diffusion* zone which is probably *about* 0·4 in. thick at a distance of 4·0 in., say, downwind from the leading edge. An exact mean value for f clearly presents a formidable problem and in the present calculation a value of 1·6 is taken by virtue of the above data, the choice of f influencing only slightly the computed theoretical rate of evaporation from eqn. (3.3).

Experimental and Theoretical Rates of Evaporation from Saturated, Plane, Rectangular Areas

Air temperature = 18·5° c, mean air speed = 1 330 ft/min, $m = 0{\cdot}15$, mean evaporating surface temperature = 17·0° c

			from eqn. (3.2)		from eqn. (3.3)	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
3·5	4·0	$3{\cdot}83 \times 10^{-5}$	$3{\cdot}89 \times 10^{-5}$	102%	$1{\cdot}94 \times 10^{-5}$	51%
5·0	4·0	4·93 ,,	5·32 ,,	109%	2·77 ,,	56%
15·0	4·0	11·02 ,,	14·17 ,,	129%	8·32 ,,	76%
5·0	2·5	3·32 ,,	3·34 ,,	101%	1·85 ,,	56%
15·0	2·5	7·39 ,,	8·84 ,,	120%	5·57 ,,	76%
2·0*	4·0	1·21 ,,			1·14 ,,	96%

(1) length (in.); (2) width (in.); (3) experimental rate of evaporation (lb/min); (4)–(7) theoretical rates of evaporation: (4), (6) lb/min, (5), (7) as percentage of (3).

* This result applies to the region between 13·0 and 15·0 in. downwind from the leading edge.

Sutton's formula (3.2) for rate of evaporation neglects the effect of lateral diffusion and consequently of the longitudinal edges, but incorporates the effect due to the leading edge. The formula (3.3) obtained by the present authors is based on a simplified problem of diffusion from an infinitely long area, with *finite width*, so that the effect of the longitudinal edges on rate of evaporation was approximated to. Both treatments make use of the concept of eddy diffusivity together with the notion that the transfer of mass is linked to that of momentum.

§ 4. CONCLUSIONS

The results given in the table show that provided the crosswind width of the evaporating area is not less than about half the downwind length, the two-dimensional solution of this limited area problem due to Sutton is found to lead to tolerable accuracy in the circumstances tested.

It is interesting to note that the experimental value for average rate of evaporation per unit length in the range 13–15 in. from the leading edge is found to exceed by only 4% the theoretical value (see table, bottom row) given by eqn. (3.3).

Finally, the results given in this paper provide further evidence to show that the approach to practical problems of turbulent diffusion and evaporation, initiated by Sutton (1949) using the concept of eddy diffusivity, is a powerful one, but in order to account theoretically for the effect on rate of evaporation of variations in length *and* width of area a full analytic, computable solution is needed to cover the effects of leading, trailing and longitudinal edges.

ACKNOWLEDGMENTS

We wish gratefully to acknowledge the generous help and advice given to us on the temperature measurements by Mr. C. Hulse and Mr. P. A. Price of the Fuel Technology Department, Sheffield University, and to express our thanks to Mr. R. G. Siddall, Mathematics Department, for assistance with the experimental and computational work involved in the preparation of fig. 3. We are also indebted to the Mining Engineering Department for the use of their wind tunnel, and to the Sheffield University Research Fund for a grant to cover the cost of apparatus.

REFERENCES

Davies, D. R., and Walters, T. S., 1951, *Quart. J. Mech. Appl. Math.*, **4**, 466.
Pasquill, F., 1943, *Proc. Roy. Soc.* A, **182**, 75.
Sutton, O. G., 1949, *Atmospheric Turbulence* (London : Methuen).
Townsend, A. A., 1951, *Proc. Camb. Phil. Soc.*, **47**, 375.

Experiments on the Mechanics of Rubber
III: Small Torsions of Stretched Prisms

By A. N. GENT and R. S. RIVLIN
British Rubber Producers' Research Association, Welwyn Garden City, Herts.

Communicated by W. P. Fletcher; *MS. received 2nd April* 1952

ABSTRACT. A theorem, relating the torsional modulus for small torsions of a rod of incompressible highly elastic material, held in a state of simple extension, to the extension ratio and force necessary to maintain the extension, is verified experimentally for rods of vulcanized rubber of rectangular cross section.

§ 1. INTRODUCTION

It has been shown (Rivlin 1948) that for an incompressible highly elastic material which is isotropic in its undeformed state, the stored energy function W must be related to the deformation through two strain invariants I_1 and I_2 which can be expressed in terms of the deformation. If a rod of such a material, of circular cross section, is subjected to a simple extension of extension ratio λ and then twisted, for sufficiently small torsions the torsional modulus is related to the extension ratio λ and the force N required to maintain the extension, by the relation (Rivlin 1949)

$$2[M/\psi]_{\psi=0}(\lambda-1/\lambda^2)=Na^2, \qquad \ldots\ldots(1)$$

where a is the radius of the rod in its undeformed state. M is the torsional couple and ψ the amount of torsion, measured per unit length of the rod in its extended state, so that $[M/\psi]_{\psi=0}$ is the torsional modulus of the rod in the limiting case when the torsion is infinitesimally small. It is seen that this relation is independent of the manner in which W depends on I_1 and I_2 for the material.

This theorem has been verified experimentally (Rivlin and Saunders 1951) using a rod of vulcanized natural rubber. Green and Shield (1951) have derived a more general theorem applicable to a uniform rod of any cross sectional form twisted about any axis parallel to its length. They obtain the result (see Appendix)

$$\frac{1}{N}\left[\frac{M}{\psi}\right]_{\psi=0}\left(\lambda-\frac{1}{\lambda^2}\right)=\frac{1}{A_0}\left[I_0-\frac{I_0-S_0}{\lambda^3}\right], \qquad \ldots\ldots(2)$$

where A_0 is the cross sectional area of the rod in its undeformed state. I_0/A_0 is the square of the radius of gyration of the unstrained cross section about the position of the axis of torsion in the undeformed body, it being assumed that this axis moves with the body when it is deformed. S_0 is the torsional rigidity about the same axis for infinitesimally small torsions of a rod having a cross section identical with that of the unextended rod, the rigidity modulus for the material being taken as unity.

In the present paper, the relation (2) is verified experimentally for the case of prisms of vulcanized rubber of rectangular cross section. For such prisms, in which the sides of the rectangular cross section have lengths a and b, I_0/A_0 is given by

$$I_0/A_0=(a^2+b^2)/12 \qquad \ldots\ldots(3)$$

and S_0/A_0 is given (Love 1927, §225) by

$$\frac{S_0}{A_0}=\tfrac{1}{3}b^2-2\,\frac{b^3}{a}\left(\frac{2}{\pi}\right)^5\sum_{n=0}^{\infty}\frac{1}{(2n+1)^5}\tanh\frac{(2n+1)\pi a}{2b}. \qquad \ldots\ldots(4)$$

§ 2. EXPERIMENTAL

The four rods used in these experiments were formed by a moulding process from a mix having the following composition in parts by weight: natural rubber (smoked sheet) 100, sulphur 2·0, zinc oxide 2·0, stearic acid 0·5, M.B.T.S. 1·0, Nonox D 0·5. Vulcanization was effected by heating at a temperature of 140° c for 45 minutes. Each of the rods was 7·46 cm in length and had steel end-pieces, by which it could be held in grips, bonded to it. Each of the rods had a rectangular cross section, the lengths of the sides a and b being given in table 1.

Table 1

Reference No. of Rod	1	2	3	4
a (cm)	1·900	1·270	0·952	0·951
b (cm)	0·477	0·478	0·478	0·950
Ratio a/b	3·980	2·654	1·992	1·001

The test-pieces were subjected to various degrees of simple extension and in each state of simple extension the relation between the torsional couple M and amount of torsion ψ for small superposed torsions was measured, as well as the force N necessary to maintain the simple extension with $\psi=0$. These measurements were carried out in an apparatus similar to that used by Rivlin and Saunders (1951) in verifying eqn. (1) experimentally. The grips and end-pieces were so designed that in each case the torsion took place about an axis passing through the centroids of the cross sections. Owing to the difficulty of measuring the torsion ψ over the central regions of the rods, where end effects would not be expected to be significant, it was measured over the whole rod (i.e. from the relative rotation of the end-pieces) and the inaccuracy involved, for various degrees of extension of the rods, was determined in a subsidiary experiment. In all cases, it was found that the error so introduced in the value of ψ was less than 1% and was therefore neglected. The extension ratio was obtained by measurements over the central region of the test-pieces where end-effects could be neglected, in the manner adopted by Rivlin and Saunders.

The measurements of corresponding values of M and ψ for various values of the extension ratio λ were carried out over a range of values of ψ (0–0·3 rad/cm) such that the resulting (M, ψ) curves were substantially linear. The values of $[M/\psi]_{\psi=0}$ were obtained as the slopes of these straight lines and are given in table 2, together with the corresponding measured values of the force N necessary to maintain the extension. For each rod the values of I_0/A_0 and S_0/A_0 were calculated from eqns. (3) and (4) respectively and the values obtained are also given in table 2. From the values of N, $[M/\psi]_{\psi=0}$ and λ, given in table 2, the corresponding values of the quantity on the left-hand side of eqn. (2), denoted α, were calculated and from the values of I_0/A_0, S_0/A_0 and λ, the values of the quantity on the right-hand side of eqn. (2), denoted β, were calculated. The values of α and β so obtained are given in table 2. It is seen that they are in reasonably good agreement.

Table 2

Rod No. 1	$I_0/A_0=0{\cdot}3198\ \text{cm}^2$,		$S_0/A_0=0{\cdot}06402\ \text{cm}^2$		
λ	1·257	1·512	1·762	2·030	2·284
N (kg)	1·906	3·11	3·97	4·68	5·31
$[M/\psi]_{\psi=0}$ (kg cm²/rad)	0·600	0·730	0·744	0·760	0·765
α	0·1965	0·2525	0·270	0·290	0·301
β	0·191	0·246	0·273	0·289	0·298
Rod No. 2	$I_0/A_0=0{\cdot}1535\ \text{cm}^2$,		$S_0/A_0=0{\cdot}05842\ \text{cm}^2$		
λ	1·264	1·524	1·784	2·036	2·294
N (kg)	1·34	2·16	2·735	3·22	3·635
$[M/\psi]_{\psi=0}$ (kg cm²/rad)	0·228	0·246	0·2485	0·246	0·246
α	0·108	0·125	0·134	0·137	0·142
β	0·106	0·127	0·137	0·142	0·146
Rod No. 3	$I_0/A_0=0{\cdot}09457\ \text{cm}^2$,		$S_0/A_0=0{\cdot}05214\ \text{cm}^2$		
λ	1·264	1·518	1·769	2·032	2·283
N (kg)	0·989	1·577	1·960	2·325	2·708
$[M/\psi]_{\psi=0}$ (kg cm²/rad)	0·1125	0·112	0·1145	0·117	0·114
α	0·0727	0·0758	0·0847	0·0900	0·0880
β	0·0736	0·0825	0·0870	0·0895	0·0910
Rod No. 4	$I_0/A_0=0{\cdot}1507\ \text{cm}^2$,		$S_0/A_0=0{\cdot}1268\ \text{cm}^2$		
λ	1·262	1·520	1·769	2·030	2·288
N (kg)	1·95	3·13	3·965	4·68	5·295
$[M/\psi]_{\psi=0}$ (kg cm²/rad)	0·421	0·412	0·392	0·375	0·358
α	0·137	0·143	0·144	0·143	0·141
β	0·139	0·144	0·1465	0·148	0·149

APPENDIX

In obtaining the result (2), Green and Shield (1951) derive a corresponding theorem for a compressible material:

$$\frac{1}{N}\left[\frac{M}{\psi}\right]_{\psi=0}=\frac{I_0-(I_0-S_0)\lambda_1^2/\lambda^2}{(1/\lambda_1^2-1/\lambda^2)A_0}, \qquad \ldots\ldots(5)$$

in which λ_1 is the extension ratio in directions perpendicular to that of the load N. Equation (2) is then obtained by writing $\lambda_1=1/\sqrt{\lambda}$ for an incompressible material. In eqns. (2) and (5), as derived by Green and Shield, M is calculated as the couple about the axis of torsion of the forces acting on the curved surface into which a plane normal to the axis of torsion is bent as a result of the torsion. This is not strictly the torsional couple required to maintain a state of torsion in the stretched prism which we have measured. The latter is defined as the moment about the axis of torsion of the forces acting on a cross section of the prism which is plane and normal to the axis of torsion in the deformed state and can be calculated by a method similar to that adopted by Green and Shield. If the prism is subjected to a small torsion superposed on a simple extension, while the surfaces of the prism parallel to the axis of torsion are force-free, then the expressions for the two couples are identical. However, if we consider that in addition to the longitudinal force a uniform hydrostatic pressure ϖ is acting on the prism, Green and Shield obtain for the couple M they have calculated

$$M=\psi\lambda_1^4\{(\Phi+\Psi\lambda_1^2)[\lambda^2 I_0-\lambda_1^2(I_0-S_0)]-\varpi(I_0-S_0)\}, \qquad \ldots\ldots(6)$$

while we obtain for the couple M which must be applied to the ends of the prism in order to maintain the state of torsion

$$M=\psi\lambda_1^4\{(\Phi+\Psi\lambda_1^2)[\lambda^2 I_0-\lambda_1^2(I_0-S_0)]\}. \qquad \ldots\ldots(7)$$

In both eqns. (6) and (7) Φ and Ψ are given by $\Phi=2\partial W/\partial I_1$ and $\Psi=2\partial W/\partial I_2$, where W is the stored-energy function for the material and I_1 and I_2 are two of the three strain invariants in terms of which it is expressed. We note that although M given by (6) involves ϖ explicitly, eqn. (7) does not. Thus, if we interpret M in the sense in which it is used in the experiments, eqns. (2) and (5) are valid whether or not a hydrostatic pressure is applied to the prism.

The apparent discrepancy between the expressions for M given in eqns. (6) and (7) may be further explained in the following manner. The presence of a hydrostatic pressure results in a uniform surface traction ϖ acting normally to the curved deformed end of the prism. This, when resolved in a plane normal to the axis of torsion, gives rise to a couple about this axis of magnitude $-\psi\lambda_1^4\varpi(I_0-S_0)$, which is exactly balanced by the couple about the axis of torsion due to the action of the hydrostatic pressure ϖ on the curved sides of the deformed prism between a curved end and a plane section normal to the axis.

ACKNOWLEDGMENT

This work forms part of a programme of fundamental research undertaken by the Board of the British Rubber Producers' Research Association. Our thanks are due to Professor A. E. Green for valuable discussions of the theoretical aspects of the problem.

REFERENCES

GREEN, A. E., and SHIELD, R. T., 1951, *Phil. Trans. Roy. Soc.* A, **244,** 47.

LOVE, A. E. H., 1927, *A Treatise on the Mathematical Theory of Elasticity,* 4th edn. (Cambridge : University Press).

RIVLIN, R. S., 1948, *Phil. Trans. Roy. Soc.* A, **241,** 379; 1949, *Ibid.*, **242,** 173.

RIVLIN, R. S., and SAUNDERS, D. W., 1951, *Phil. Trans. Roy. Soc.* A, **243,** 251.

LETTERS TO THE EDITOR

Transparent Conducting Films

Research on thin electrically conducting films, in connection with electrically heated windows, has led to the following observations on gold films :

1. For very thin (e.g. 60Å) gold films sputtered directly on glass, the resistivity is much higher than that of the metal in bulk. Also, the optical absorption is markedly higher than that calculated from the bulk optical constants. The transmitted colour is variable from blue to blue-green. On heating the film, the colour becomes first definitely blue and then, at a higher temperature, pink. There is evidence of grain growth during these changes. There are, however, theoretical grounds for expecting the transmitted colour of such a thin film to be pale yellow, if the metal were in the form of a uniform homogeneous layer.

2. The above-mentioned properties are markedly influenced by the process used for cleaning the glass. Cleaning the glass with dry precipitated chalk tends to impart to the films a low resistance, low absorption and a colour inclining toward ' straw '. Chemical cleaning, on the other hand, tends to give the films a high resistance, high absorption, and a consistent blue colour.

3. The effect of other substrates is still more marked. Oxide substrates have been studied in particular. On a substrate of sputtered zinc oxide the gold film is invariably blue, and has an exceptionally high resistance. On sputtered bismuth oxide, however, the film is of a pale straw colour, while the resistance and optical absorption are consistently low. In this case, subsequent heating for a few minutes at about 200° C results in still lower values of resistance and absorption, not far from those based on the bulk constants (with a ' thinness ' correction).

Observations 2 and 3 appear to be novel, and the practical applications of 3 have been patented. The following are typical results for sputtered gold films of thickness 60–70Å (on a weight-bulk density basis). The optical results are overall figures including the effects of the thin glass support, and of the oxide film. The transmission factor T, and the absorption factor α (for light incident on the coated side of the glass support) refer to tungsten filament light, colour temperature about 2 700° K, and R is the resistance between two opposite edges of a square of film.

Substrate	R	α	T
Chalk-cleaned glass	60–70	24	62
100 Å bismuth oxide, on glass, with final heating	10	13	75

It is suggested that certain substrates, like bismuth oxide, strongly favour adhesion of the gold, with low mobility over the surface and a minimum of aggregation, resulting in good homogeneity of the films, at least up to a certain thickness. Other substrates favour the opposite effect. In any case, the results point to a new field for the study of surface forces, and of the properties of gold (and perhaps other) films in relation to their structure and to the bulk constants of the metal. In the present work structural and other observations are being extended. The sputtered films were produced by the method described by Preston (1950).

The work described above has been carried out in the Light Division of the National Physical Laboratory on behalf of the Ministry of Supply. This letter is published by permission of the Director of the Laboratory and with the approval of the Chief Scientist, Ministry of Supply.

National Physical Laboratory,
Teddington, Middlesex.
5th June 1952.

E. J. Gillham.
J. S. Preston.

Preston, J. S., 1950, *Proc. Roy. Soc.* A, **202**, 449.

Area Contacts on Germanium

The behaviour of point contacts on germanium and of p–n junctions within single crystals of the material has been extensively studied by many authors, but the properties of metallic contacts of large area have received comparatively little attention. Some of these properties are described below. The observations refer mostly to gold contacts applied by evaporation in vacuum to n-type germanium of 5 ohm-cm resistivity and a Hall coefficient of about 10^4 cm^3/coulomb. They apply qualitatively also to specimens of higher conductivity. The germanium surfaces were etched before application of the electrodes.

The voltage–current relations of contacts prepared in this way are qualitatively similar to those obtained with point contacts. The reverse currents usually show a partial saturation at small reverse voltages. The saturation currents which can be evaluated by a simple geometrical extrapolation are found to be temperature dependent in accordance with an equation of the form $i_s = i_0 \exp(-\phi/kT)$, where ϕ varies from contact to contact between 0·60 and 0·72 ev. No significant differences are obtained with contacts of gold, silver and copper and no improvement of the rectifying properties could be achieved by forming processes of the type normally employed for point contacts. This may be due to the fact that equivalent current densities cannot be obtained on area contacts without causing burn-out.

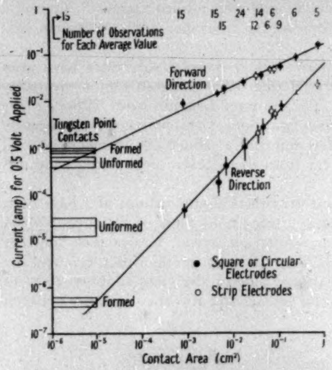


Fig. 1. Average reverse and forward currents as functions of electrode area for 0·5 volt applied to the assembly as a whole.

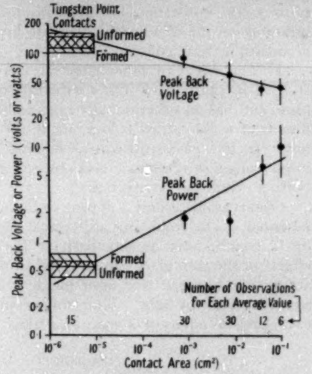


Fig. 2. Average peak back voltage and power as functions of electrode area.

The slope (45°) of the lower line on fig. 1 shows that the total current for a small reverse voltage is directly proportional to the contact area. This holds for electrodes of different shapes and implies that the current density is uniform (no edge effects). The differential conductance in the saturation region is also found to be a linear function of contact area. In the forward direction the series resistance of the bulk germanium becomes important and the total voltage applied to the specimen as a whole is then no longer equal to the barrier voltage. Forward currents at a given total voltage are not directly proportional to the contact area but increase less rapidly, as shown in the figure. This is due to geometrical factors and to the effect of injected carriers on the resistance in series with the contact barrier. For small electrodes the series resistance is essentially a *spreading* resistance, and the injected carriers which reduce the resistivity of the material near the contact are effective in the most important region. This accounts for the diminution of the rectification ratio with increasing contact area, and such an effect must be expected in all materials in which carrier injection occurs.

Figure 2 shows the variation of peak back voltage and peak back power with electrode area. In the neighbourhood of the voltage turnover the contact temperature can no longer be assumed to be equal to the ambient temperature. There is, however, some evidence (e.g. from experiments on electrodes of different shapes and different conditions of heat dissipation) that the turnover phenomenon is not solely due to contact heating. The peak back voltage and the peak back power depend on the ambient temperature in a manner very similar to that of point contacts on germanium and lead sulphide.

For purposes of comparison, some results for formed and unformed tungsten point contacts have also been included in the figures. The forming process has little effect on the peak back conditions, but it sensitively affects the magnitude of reverse currents for small voltages. According to fig. 1 the behaviour of the (unformed) evaporated contacts is similar to that of the *formed* point contacts, a fact which must have an important bearing on the mechanism of the forming process and which is not yet adequately understood.

Thanks are due to Professor R. W. Ditchburn for his interest in this work.

Department of Physics,
University of Reading.
5th June 1952.

J. W. Granville.
H. K. Henisch.

Contact Properties of p–Type Germanium

The interpretation of many phenomena observed on contacts between n-type germanium and various metals or between *two* specimens of n-type germanium involves specific assumptions concerning the existence and position of surface levels. Pfann and Scaff (1949, 1950) and Fan and Bray (1949) have already drawn attention to the fact that high-voltage rectification and transistor action can be observed on p-type as well as n-type germanium. However, the interpretation of these effects on p-type material demands quite different assumptions concerning surface levels. These problems still await a satisfactory solution. Meanwhile, a more detailed comparison between the contact properties of n- and p-type germanium seemed desirable and some results are described below.

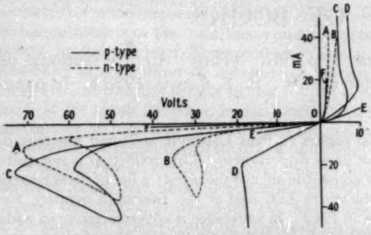


Fig. 1. Voltage–current relations for point contacts on germanium: A, n-type starting material etched; B, n-type starting material after polishing; C, p-type material produced by heat shock (etched); D, p-type material as for C after polishing; E, p-type characteristic given by Torrey and Whitmer; F, p-type characteristic of a formed contact given by Pfann and Scaff. (Curves A–D, which refer to *unformed* contacts, were obtained by oscillographic methods. The close similarity between the peak back voltages of curves A and C is fortuitous.)

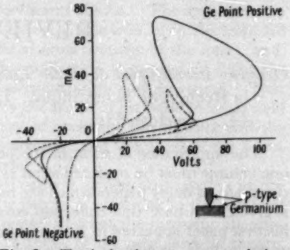


Fig. 2. Typical voltage–current relations for contacts between two specimens of p-type germanium (oscillographic method).

Figure 1 shows typical characteristics of an unformed tungsten point contact on etched p-type germanium and also on a polished surface. The material was prepared from n-type germanium by heat shock (heating to about 800° c, followed by rapid cooling). Typical characteristics observed on the starting material are also shown, as well as p-type curves given by Torrey and Whitmer (1948) and (for a *formed* contact) by Pfann and Scaff (1950). For comparison, n- and p-type rectification curves have been drawn in the same quadrants. The peak back voltages measured on the present specimens vary a good deal from point to point, more so than on the n-type starting material. For a given point, the peak back voltage and power decrease rapidly and continuously with increasing temperature up to 100° c. Anomalous thermo-electric effects of the type observed by Granville and Hogarth (1951) on polished surfaces of n-type material were *not* observed on the present p-type specimens. As for n-type germanium, there are no significant differences between the average peak back voltages observed for different whisker materials. Nickel, platinum, tungsten and copper gave substantially similar results.

Figure 2 gives voltage–current relations obtained for four different contacts between two specimens of p-type material. These relations show that a barrier of considerable height must exist at these junctions.

Thanks are due to Professor R. W. Ditchburn for the support of this research.

Department of Physics,
University of Reading.
26th May 1952.

J. W. Granville.
H. K. Henisch.
P. M. Tipple.

Fan, H. Y., and Bray, R., 1949, *Phys. Rev.*, **76**, 458.
Granville, J. W., and Hogarth, C. A., 1951, *Proc. Phys. Soc.* B, **64**, 488.
Pfann, W. G., and Scaff, J. H., 1949, *Phys. Rev.*, **76**, 459; 1950, *Proc. Inst. Radio Engrs., N.Y.*, **38**, 1151.
Torrey, H. C., and Whitmer, C. A., 1948, *Crystal Rectifiers* (New York : McGraw-Hill).

REVIEWS OF BOOKS

Oeuvres scientifiques de Paul Langevin. Pp. 687. (Paris : Centre National de la Recherche Scientifique; London agents : H. K. Lewis, 1950.) In paper 44*s.* ; in cloth 53*s.*

The appearance of most of the original papers of Paul Langevin collected together in one volume must be regarded as an event of great importance. This collection is not quite complete, as many papers which are readily accessible are not included; a list of these papers is given. In time this collection covers the entire first half of the twentieth century, for his first paper appeared in 1902 and his last in 1949 : in substance it covers a range so wide as to include important papers on ionization of gases, kinetic theory, thermodynamics, electromagnetic theory, magnetism, relativity, radioactivity, ultrasonics, nuclear physics and gravitation. This is indeed a contribution to physics in the grand manner. An obituary notice of Langevin has been published in the *Proceedings of the Physical Society* (1947, **59**, 1041), where tribute to his immense range of achievement was made. His earlier work, coming as it did at a vital period in the development of physics, played a most important role in effecting the transition from the classical to the modern point of view, and can be regarded as forming a bridge between the statistical mechanics of the late 19th century and the early quantum ideas of the 20th century. In particular, Langevin's work is an excellent example of the power and scope of classical statistics in dealing with the fundamental collisional phenomena involved in the conduction of electricity through gases. Then, again, it is interesting to note that he was a pioneer worker in relativity, and also that Einstein first learned of de Broglie's original notions of wave mechanics from Langevin. For these reasons this volume of collected papers is, and probably will continue to be, of profound interest to the student of physics, and particularly to the young research worker, since, in addition to important theoretical papers, it also gives numerous papers describing his experiments.

Langevin's first papers dealt with the properties and motions of ions (the existence of free electrons in gases had not yet been demonstrated), and his interest in this subject no doubt originated during the time he spent at the Cavendish Laboratory under Sir J. J. Thomson. An outstanding problem which excited the interest of physicists at the turn of the century was that of the explanation of the electrical conductivity produced by x-rays or radioactive emanations. It was the application of classical kinetic theory to the experimental determinations of conductivity, and of current decay, which enabled the nature of the ion carriers to be found, the processes of ionization and recombination to be investigated, and the mechanisms of ionic transport, diffusion and recombination to be understood in terms of the collisions of elastic sphere atom models.

By consideration of the polarization of neighbouring neutral atoms produced by a charged ion, and also the resulting attraction, effecting reduction of the free path, the consequent collection of molecules to form large clusters about a single ion was envisaged. These charged molecular groups, whether singly or in clusters, were shown to be the carriers of electricity. It was in this particular field that Langevin made his first outstanding contribution. In the series of papers in 1902–1905 (*Ann. Chim. Phys.*, **5**, 245) culminating in the famous paper *Une formule fondamentale de théorie cinétique*, he treated the motion of ions in gases, and rigorously derived the well-known formula for the drift velocity W:

$$W=0{\cdot}815(Ze/m)(\lambda/C)\{(m+M)/M\}^{1/2},$$

of an ion of charge e, free path λ, r.m.s. velocity of agitation $C(\gg W)$, mass m moving under a field Z in a gas of molecular mass M in which the ion exerts no electric forces except at collision (low gas pressure). The ions are considered to be nearly in statistical equilibrium with the gas atoms: thus restriction is now known to be unnecessary, since the formula also applies to the motions of electrons under a regime of elastic conditions when the electrons are not in statistical equilibrium with the gas atoms, provided, however, that their energy distribution is Maxwellian. This latter consideration, again, only affects the numerical coefficient, which only varies from about 2/3 to unity over a wide range of electron energy distribution functions. Replacement of the force Ze by that due to the concentration gradient leads to Langevin's formula for the diffusion coefficient of ions. The application of his formulae to the measurements of the mobilities of clusters, single ions, or of electrons led to determinations of the mass factor and of the collisional cross sections of the ions: and, in fact, in the hands of Townsend and Tizard, such application suggested in 1913 the possibility of a variation of atomic cross section in collisions with slow electrons. That Langevin had peculiar insight in this subject is seen by the fact that his basic formula held the field for about 25 years, in fact until the discovery of exchange phenomena and the calculations of Mott and Massey based on quantum scattering, and was used in the interpretation of the experimental work of Tyndall and Powell.

The other great contribution with which Langevin's name is always associated is in the theory of magnetism, and this was made in 1905. The theory of electrons in the hands of Lorentz was making great strides, but doubt was expressed whether the magnetic theories of Ampère and Weber could be expressed adequately in terms of the new electron theory. It was Langevin who set these doubts at rest. Curie's work on the temperature dependence of paramagnetism was almost completely explained by Langevin on the basis of the classical kinetic theory on the assumption that paramagnetic susceptibility was due to the reserve of permanent magnetic dipoles in each molecule. The results deduced that over a large range of temperature paramagnetic susceptibility is inversely proportional to the absolute temperatures—known as the Curie–Langevin law. This work today still forms a most valuable introduction to the electron theory of magnetism. The full explanation, however, had to await the application of the quantum theory of spin, with its restrictions on the orientations of the molecular magnets in a magnetic field.

This book of 687 pages gives a *curriculum vitae* of Langevin and a complete bibliography of his scientific works, and has altogether thirteen chapters. In spite of its size, this volume is quite pleasant to handle, and the print is clear and easy to read. The first chapter describes his experimental works on ionization of gases by x-rays, his experimental and theoretical work on ionic mobility and recombination, including the polarization of neutral atoms produced by the ion charge, and the properties of secondary x-rays; he contributed to these subjects throughout his life. The second chapter is devoted to considerations of ionization in the atmosphere and the properties of ion clusters.

The next three chapters are of a theoretical character and give his papers on fundamental kinetic theory, electromagnetic inertia, relativity, and also his classic work in the theory of magnetism. The sixth chapter presents a collection of his papers on relativity, and this is followed by chapters on radioactivity and units. Then follow some twenty pages devoted to a critical review of classical, quantum and statistical mechanics. Chapters ten and eleven are concerned with his contribution to ultrasonics, under-water acoustics and divers technical problems : these represent his scientific war work in World War I. He originated and developed the piezoelectric under-water detector, and carried out trials in Toulon harbour. The two final chapters are devoted to his most recent work on nuclear physics and gravitation, and are of a fundamental nature.

It is not possible within the scope of a short review to discuss the chapters in detail : it is sufficient to stress the scope and quality and importance of the work described, to admire the clarity of its exposition, and very strongly to recommend this collection to all interested in the history and development of an important aspect of modern physics.

F. LLEWELLYN JONES.

CONTENTS FOR SECTION A

PAGE

Prof. L. F. Bates. Some Post-War Developments in Magnetism (Presidential Address) . . . 577

Dr. D. M. S. Bagguley and Dr. J. H. E. Griffiths. Paramagnetic Resonance and Spin–Lattice Relaxation Times in the Copper Tutton Salts . . . 594

Dr. P. T. Landsberg. A Note on the Theory of Semiconductors . . . 604

Dr. Suraj N. Gupta. Quantization of Einstein's Gravitational Field : General Treatment . . . 608

Mr. H. N. V. Temperley. A New Theory of Liquid Helium . . . 619

Mr. D. Welti and Dr. R. F. Barrow. The Absorption Spectra of Gallium and Indium Monofluorides . . . 629

Mr. D. S. Schonland. On the Utilization of Multiple Scattering Measurements . 640

Mr. H. W. Wilson and Mr. G. M. Lewis. Decay Schemes of ^{151}Sm and ^{155}Eu . 656

Mr. A. J. M. Hitchcock. The structure of Lithium 6 . . . 661

Dr. A. Dalgarno. The Photo-Ionization Cross Section of Methane . . . 663

Letters to the Editor :

Dr. B. Bleaney and Mr. K. D. Bowers. The Cupric Ion in a Trigonal Crystalline Electric Field . . . 667

Mr. P. A. Lee, Prof. E. A. Stewardson and Mr. J. E. Wilson. The M_V Absorption Edges of Tm and Er and the Origin of the Er M_V Absorption Lines . . . 668

Dr. E. A. Power. The Nature of the Neutral π-Meson . . . 670

Mr. F. K. Goward and Mr. J. J. Wilkins. Multiple Resonances in Photo-disintegration Cross Sections . . . 671

Dr. H. N. Yadav. The Scattering of Positrons by Nuclei . . . 672

Mr. R. Batchelor. Neutron Energy Measurements with a Helium 3 Filled Proportional Counter . . . 674

Mr. G. A. Dissanaike and Mr. J. O. Newton. The decay of the 2·4 Mev Level in ^{9}Be . . . 675

Mr. K. Maack Bisgård. Internal Conversion Electrons from the ^{233}U decay 677

Reviews of Books . . . 678

Contents for Section B . . . 685

Abstracts for Section B . . . 686

ABSTRACTS FOR SECTION A

Some Post-War Developments in Magnetism, by L. F. Bates.

ABSTRACT. A survey is given of some of the more outstanding developments in magnetism during the post-war years, including the great extension of our knowledge of the ferromagnetic domain through the theoretical work of Néel and others and by the systematic use of the Bitter powder figure techniques; the main ferromagnetic, paramagnetic and nuclear resonance phenomena and experiments on the diffraction of neutrons by anti-ferromagnetic crystals are also described.

Paramagnetic Resonance and Spin–Lattice Relaxation Times in the Copper Tutton Salts, by D. M. S. Bagguley and J. H. E. Griffiths.

ABSTRACT. Paramagnetic resonance absorption in single crystals of the copper Tutton salts has been investigated at 3·10 cm wavelength between 90° K and room temperature. In salts diluted with zinc Tutton salt the hyperfine lines are clearly resolved at 90° K but broaden and merge into one another as the temperature is raised. From these measurements the spin–lattice relaxation time τ can be estimated and is found to be about 2×10^{-9} sec. It is found that $1/\tau$ varies linearly with temperature between 150° K and 290° K. An investigation of the crystalline field in the Tutton salts shows that for the concentrated salts the field has a strong rhombic component above 150° K, whilst in the salts diluted with zinc to less than 1% copper the crystalline field has tetragonal symmetry at 290° K, but the tetragonal axis lies in a plane perpendicular to that containing the tetragonal axis at 90° K.

A Note on the Theory of Semiconductors, by P. T. Landsberg.

ABSTRACT. One obtains two different theories of semiconductors, depending on whether the free energy due to the spin of electrons in donor (or acceptor) centres is taken into account (approach A) or not (approach B). It is shown (i) that the former approach can be carried through for a degenerate semiconductor, (ii) that it applies *formally* to the case when the interaction between impurity centres and conduction electrons is taken into account, provided one of the symbols is reinterpreted, and (iii) that it is to be preferred. Approach B has, however, been used almost universally. It is shown that it corresponds to the unlikely situation in which fifty per cent of the electrons from each impurity level come from paired states, and the remaining fifty per cent from an unpaired state having the same activation energy.

Quantization of Einstein's Gravitational Field: II—General Treatment, by Suraj N. Gupta.

ABSTRACT. The quantization of the complete gravitational field is carried out by extending the work of an earlier paper. The main obstacles in the quantization of Einstein's field are overcome by expressing the field quantities in the Riemannian space as expansions in the flat space, and then splitting the gravitational field into the linear and the non-linear parts. The linear part of the gravitational field is regarded as the free gravitational field, while the non-linearity is treated as a direct interaction between the gravitons. This treatment is quite general, but it suffers from the usual limitations of the perturbation method.

The gravitational self-energies of the photon and the electron are also investigated by using the linear approximation of the gravitational field. It is found that the photon self-energy vanishes unambiguously, while the electron self-energy is quadratically divergent.

A New Theory of Liquid Helium, by H. N. V. Temperley.

ABSTRACT. A model of liquid helium is described, which incorporates the main features of the Landau (1941) and London (1939) theories, making use also of Mayer's concept of clusters. The ordinary process of liquefaction is associated with the formation of clusters in coordinate space, while the λ-transition is probably associated with the formation of small clusters in momentum space. The latter process is probably prevented, or greatly modified, in helium 3.

Possible reasons for the appearance of superfluid effects are given, and it is suggested that the unique properties of helium II may be traced to the fact that the interactions between two atoms is of about the strength needed to form a molecule.

The Absorption Spectra of Gallium and Indium Monofluorides, by D. Welti and R. F. Barrow.

ABSTRACT. Gaseous GaF and InF are formed when the metals are heated with aluminium trifluoride, and sufficient pressures of the monofluorides for observation of their absorption spectra are obtained at about 1 000° c for GaF and 1 400° c for InF. The three ultra-violet band-systems, $^1\Pi$, $^3\Pi_1$, $^3\Pi_{0^+} \leftarrow {}^1\Sigma^+$ expected by analogy with GaCl and InCl have been photographed on spectrographs of moderate dispersion, and vibrational analyses have been carried out. The $^1\Pi$ states converge rapidly to give rather sharp dissociation limits: the products are probably Ga or In, $(^2P_{3/2})$, $+F(^2P_{1/2})$. The bearing of the present results upon the interpretation of the electronic levels of TlF and on the dissociation energy of AlF is briefly considered.

On the Utilization of Multiple Scattering Measurements, by D. S. Schonland.

ABSTRACT. The determination of the product pv for particles leaving tracks in photographic emulsions is discussed on the basis of measurements of the scattering in the track made by the 'coordinate' method (Fowler 1950), allowing for errors of measurement due to grain size. The use of Fowler's method with the comparatively large interval lengths used in practice is shown to be justifiable in spite of its neglect of much information contained in the track. (This neglect increases the root mean square error in the estimation of the scattering over the minimum attainable error by a factor less than approximately 2·6.) A possible method of estimating pv, which approximates to the 'best possible', based on the use of mean squares, is considered, and the root mean square relative error in the estimation found, the necessity for cutting off the scattering distribution at large angles being taken into account. A possible method of estimation based on the use of mean moduli is also discussed.

Decay Schemes of ^{151}Sm and ^{155}Eu, by H. W. Wilson and G. M. Lewis.

ABSTRACT. Examination of the relatively long-lived isotopes produced by pile irradiation of samarium using the proportional tube spectrometer leads to the elucidation of the decay schemes of ^{151}Sm and ^{155}Eu. The conclusions are supported by experiments involving β–e^- and β–γ coincidences, utilizing Geiger–Müller and scintillation counters.

The Structure of Lithium 6, by A. J. M. Hitchcock.

ABSTRACT. The energies of the lowest states of ^{6}Li have been calculated from the standpoint of the shell model, using recent data for the proton–neutron forces. It is concluded that it is necessary to take into account the tensor force in discussing the lowest states of this and other odd–odd nuclei.

The Photo–Ionization Cross Section of Methane, by A. Dalgarno.

ABSTRACT. The photo-ionization cross section of methane is calculated using self-consistent wave functions. The value at the spectral head (105 850 cm^{-1}) is found to be $9{\cdot}4 \times 10^{-17}$ cm^2. An approximate general formula is employed to obtain results in a wave number range up to 200 000 cm^{-1} and the associated oscillator strength is computed. It is pointed out that exchange effects probably reduce the cross section. With the aid of the general formula an estimate is also made of the photo-ionization cross section of argon.

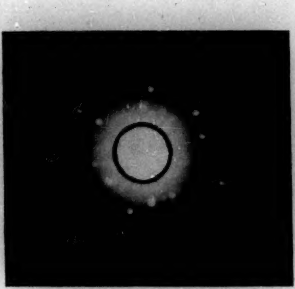


Fig. 1. Laue photograph. Tin crystal grown in graphite.

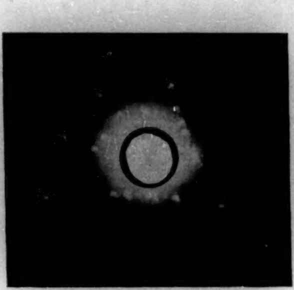
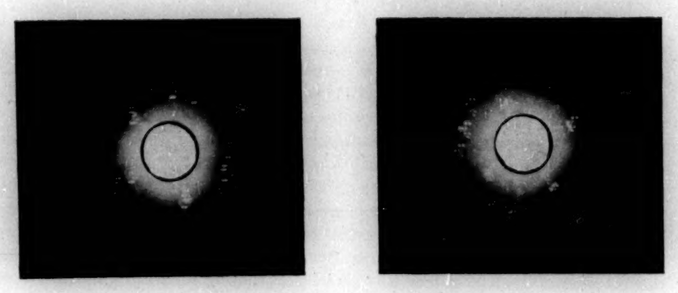


Fig. 2. Tin crystal grown in sindanyo.



Figs. 3 and 4. Crystals grown in glass.